PHOTOCOUPLED TIME-RESOLVED SMALL ANGLE X-RAY SCATTERING APPLIED TO A BLUE LIGHT SENSING PROTEIN

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Jessica Scherrer Lamb

January 2009

© 2009 Jessica Scherrer Lamb

ALL RIGHTS RESERVED

PHOTOCOUPLED TIME-RESOLVED SMALL ANGLE X-RAY SCATTERING APPLIED TO A BLUE LIGHT SENSING PROTEIN

Jessica Scherrer Lamb, Ph.D.

Cornell University 2009

Many molecules relevant to life respond to changes in their environment by large changes in shape, by which the carry out an important function. Understanding of these systems can be gained through study with methods which can map out the global conformational changes the molecule undergoes over time. Time-resolved small angle x-ray scattering is one such method. We have engineered a photocoupled flow cell which allows us to initiate changes in proteins via light excitation and measure them with a capillary focused x-ray beam. We have also considered the analysis of time-resolved SAXS data, which can be acquired using such a system, with recently popular reconstruction methods. These techniques have been applied to the study of VVD, a blue light photosensor in the PAS-LOV family of proteins. This large family of proteins is found in all kingdoms of life and is important for transducing environmental signals into biological actions. These studies provide insight into rapid dimerization of the protein upon photoexcitation. Such information is important to understanding not only VVD signaling but also the mechanisms by which other PAS-LOV proteins operate.

BIOGRAPHICAL SKETCH

Jessica Lamb lived her first year of life in the college town of Bethany, WV. Her parents brought her to Washington, CT, and then to Pennsburg, PA where she graduated from Perkiomen School. In addition to pursuing academic achievements she enjoyed many extracurricular activities in the performing arts. To this day, she is thankful to be able to speak in front of others without her knees wobbling.

Jessica attended Dartmouth College from 1998-2002. The first day she arrived, she met Alex Glocer, who she would later marry. She also interned in the lab of Professor Nunes and learned to enjoy the rewards and trials of making an experiment work. After college, she moved to Ithaca, New York to attend graduate school. In 2003 she joined the Pollack lab and began to truly pursue a career in biophysics.

This document is dedicated to all the many wonderful teachers who impacted my life over the years, from Mrs. Cohen in third grade, who when I said I didn't want to be a scientist asked me "Why not?", to my parents, who in addition to lessons on Jane Austen and the *Web of Life* taught me to be an independent woman.

ACKNOWLEDGEMENTS

I would like to acknowledge and thank my advisor Lois Pollack for her guidance, support, and enthusiasm throughout my graduate career. I would also like to thank all past and present members of the Pollack Group, including Kurt Andresen, Lisa Kwok, Heather Smith, Hye Yoon Park, Xiangyun Qiu, Suzette Pabit, Li Li and Steve Meisburger for their collaboration, insight, and friendship.

This work would not have been possible without the help of many members of the CHESS staff. Particular thanks to Sterling Cornaby, Ken Finklestein, Detlef Smilgies, Don Bilderback, and especially Arthur Woll for their expertise and assistance. I would like to thank Brian Crane and Brian Zoltowski for their insight into the workings of VVD and for many batches of protein. I would also like to acknowledge and thank Gary Whitaker, Shoshannah Roth, and Xiangjie Sun for providing HA samples.

Several facilities were used in the course of this research. Some of the computations described were performed on the cluster operated by the Cornell Center for Material Science. SAXS data were collected at the Cornell High Energy Synchrotron Source and the Advanced Photon Source. Fabrication of certain components was done at the Cornell Nanofabrication Facility.

I would like to thank Nozomi Ando and Gil Toombes for occasional help with SAXS-related problems, and for friendship. I would also like to thank Thalia Mills, Ally King, Eric Ryan and Robin Sampson for making my years in Ithaca fun and interesting.

Mom and Dad, thank you for your support and love throughout the years. I would not be here without you. Megan, you're a great sister; you can always make me laugh. And Alex, you have given me more than I can say. I love you all.

This work has been supported by the Nanobiotechnology Center and the National Institutes of Health Program Project Grant on RNA folding. It has also been indirectly supported by the funding agencies for the facilities listed above.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Biomolecular shapes and dynamics

On a physical level, all living organisms are composed of many specific and highly adapted molecules. Most of these carry out necessary biological functions to contribute to the proper function of the organism as a whole. Like a key must fit a lock in order to do its job and open the door for which it was made, these biological molecules need to acquire specific shapes in order to properly function. When formed, biomolecules carry out a wide array of tasks. They might splice themselves out of genes[1] or add phosphate to other molecules[2]. Some transport molecules throughout the body, others carry molecules across cells[3]. Biomolecules absorb light for information[4] or for food[5].

Biomolecular conformations are important, but molecular dynamics may be even more important. For a person to even lift a finger requires the contraction of muscles, which in turn is based on the rotation of myosin, a ubiquitous structural protein[6], as well as movements and changes in many other molecules. Knowledge of changes in molecular structure is critical to understanding function, and in turn, how to regulate or mimic them to prevent and treat disease and enhance biological function. In this work we will discuss investigations of conformational changes in a number of biomolecules. Discoveries about these varied molecules contribute to understanding them and the families to which they belong.

Translation of sensory information into biological action is an important

function enabling organisms to respond to their environment. The PAS (Per ARNT Sim) protein superfamily is characterized by a domain which couples environmental stimuli to molecular action. This domain consists of a particular protein fold which is found in all kingdoms of life[7]. Our studies focus on the PAS-LOV (light-oxygen-voltage) protein Vivid. Vivid is a signaling protein involved in the circadian clock regulation of the fungus *neurospora crassa*[8]. While the mechanism for such regulation remains to be elucidated, it is clear that it must involve a conformational change[9].

Many viruses attack cells by means of surface proteins which fuse the viral capsid to the cell membrane, enabling the invader to empty its genetic contents into the target cell. The action of some of these proteins is triggered by a pH jump. The most well studied of these, influenza hemagglutinin, is known to fall into to this category[10]. It undergoes a large structural rearrangement upon acidification of the endosome. However, the complete post-fusion conformation is still unknown, as is the precise cause of membrane fusion.

Interest in the function of RNA is growing as awareness of its biological importance increases. RNA, like DNA, has the ability to store genetic information[11] but can also play a catalytic role[1]. There is interest in how it goes from a molecule based largely on base-paired secondary structure, like DNA, to a molecule with a complex tertiary fold which enables catalytic activity. The folding of self-splicing introns is an example of this phenomenon.

Many techniques are useful for imaging molecular structure. Electron microscopy is popular for large molecules[12;13], particularly due to enhancements in surface reconstruction techniques[14]. Atomic force microscopes have been used to image DNA[15] and other molecules. Crystallography is a widespread

technique that uses X-rays to obtain precise measurement of molecular structure, sometimes to resolution as good as 1 Å. However, these techniques are limited in application to dynamic situations.

Optical techniques can be used to collect time-resolved data, but they often resolve only local structural changes in a molecule. A good complement to all of these techniques is small angle X-ray scattering (SAXS). Unlike most optical measurements, SAXS probes global molecular structures. Because the measurement is carried out in solution, many methods typically used to time-resolve optical measurements are compatible with SAXS. It also allows the measurement of biologically relevant states which are difficult to crystallize.

1.2 Scope of dissertation

In addition to scientific investigations into the workings of the varied molecules mentioned above, we will also be discussing SAXS as a technique which can be applied to these situations. Many methods for time-resolving molecular changes in response to various stimuli are presented. Analysis, and the use of somewhat new reconstruction methods to obtain low-resolution three dimensional models from SAXS are also discussed. These techniques are used to increase the understanding of a number of important biological systems.

This document begins with a brief introduction to the theory and uses of SAXS, and goes on to discuss its application to time-resolved problems. Several methods for making time-resolved SAXS measurements are mentioned. In particular we outline a new technique that uses a laser to initiate the molecular reaction in a flow cell.

To this end we have explored the use of focusing capillaries in SAXS measurements. These tools aid in studies involving flow cells because of their ability to concentrate the necessary X-ray intensity into a small spot size. This increases the precision of time-resolved measurements in these devices where time is correlated with spatial position.

Analysis methods are important to any measurement. In the past several years it has become popular to use computer programs to reconstruct a molecular shape envelope from the SAXS curve. To our knowledge, these reconstructions had not previously been applied to time-resolved data. We tested their utility for this application and discussed the pros and cons of the approach.

Using these methods we carried out experiments on the PAS-LOV protein VVD mentioned above. We used the photoexcitation flow cell to trigger conformational changes in this blue light absorbing protein. From these measurements we were able to determine the minimum time scale for important changes in the protein. We also collected data on the light-excited dimer and conformations important to the association of this protein. VVD is of particular interest because it is a member of the varied yet significant PAS-LOV family.

REFERENCES

1. Kruger, K., P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech. "Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena." *Cell*, **31**(1): 147–157, 1982.

2. Huala, E. "*Arabidopsis* NPH1: A protein kinase with a putative redox-sensing domain." *Science*, **278**(5346): 2120–2123, 1997.

3. Vale, R. "The molecular motor toolbox for intracellular transport." *Cell*, **112**(4): 467–480, 2003.

4. Briggs, W. and E. Huala. "Blue-light photoreceptors in higher plants." *Ann. Rev. Cell Devel. Biol.*, 1999.

5. Green, B. and D. Durnford. "The chlorophyll-carotenoid proteins of oxygenic photosynthesis." *Annu. Rev. Plant. Physiol. Plant. Mol. Biol.*, **47**: 685–714, 1996.

6. Baker, J. E., I. Brust-Mascher, S. Ramachandran, L. E. LaConte, and D. D. Thomas. "A large and distinct rotation of the myosin light chain domain occurs upon muscle contraction." *Proc. Natl. Acad. Sci. U.S.A.*, **95**(6): 2944–9, 1998.

7. Taylor, B. and I. Zhulin. "PAS domains: internal sensors of oxygen, redox potential, and light." *Microbiol. Mol. Biol. R.*, **63**(2): 479, 1999.

8. Loros, J. J. and J. C. Dunlap. "Genetic and molecular analysis of circadian rhythms in Neurospora." *Annu. Rev. Physiol.*, **63**: 757–94, 2001.

9. Zoltowski, B. D., C. Schwerdtfeger, J. Widom, J. J. Loros, A. M. Bilwes, J. C. Dunlap, and B. R. Crane. "Conformational switching in the fungal light sensor Vivid." *Science*, **316**(5827): 1054–7, 2007.

10. Earp, L., S. Delos, H. Park, and J. White. "The many mechanisms of viral membrane fusion proteins." *Curr. Top. Microbiol. Immunol*, **285**: 25–66, 2005.

11. Lehninger, A. L., D. L. Nelson, and M. M. Cox. *Lehninger Principles of Biochemistry*. W. H. Freeman & Co., 4th edition, 2005.

12. Böttcher, C., K. Ludwig, A. Herrmann, and M. van Heel. "Structure of influenza haemagglutinin at neutral and at fusogenic pH by electron cryomicroscopy." *FEBS Lett.*, 1999.

13. Li, F., M. Berardi, W. Li, M. Farzan, and P. Dormitzer. "Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain." *J. Virol.*, **80**(14): 6794–6800, 2006.

14. Subramaniam, S. and J. L. S. Milne. "Three-dimensional electron microscopy at molecular resolution." *Ann. Rev. Biophys. Biomol. Struct.*, **33**: 141–55, 2004.

15. Mou, J., D. M. Czajkowsky, Y. Zhang, and Z. Shao. "High-resolution atomic-force microscopy of DNA: the pitch of the double helix." *FEBS Lett.*, **371**(3): 279–82, 1995.

CHAPTER 2

SMALL ANGLE X-RAY SCATTERING

2.1 Basic scattering theory

Small angle X-ray scattering (SAXS) derives information from the pattern of X-rays scattered by molecules at low angles. Scattering data originates from the interaction of X-rays with electrons in the object under consideration through Thompson scattering[1]. These coherently scattered X-rays add according to their origins, as illustrated in Figure 2.1. Constructive interference will produce a high intensity pattern of X-rays in the forward scattering angles, with smaller objects scattering out to further angles. The result is data characteristic of the size and shape of the object being probed. Through analysis the experimenter can gather valuable, albeit low-resolution information about the size and general shape characteristics of an object.

This method is useful for studying the structure of biomolecules, such as proteins and ribozymes, in aqueous solution. The electron densities of such molecules are greater than those of their solvents so that the latter can be treated as background, and a scattering signal can be measured from the molecules alone. If they are dispersed randomly in solution, the scattering from the individual molecules will be additive. This increases the total scatter, but also averages signals from different orientations. The data is therefore a radially symmetric pattern of scattering intensity versus scattering angle, 2θ.

The general equation for the scattering from a dilute, finite macromolecule,

$$I(q) = V \int_0^\infty 4\pi r^2 \, \mathrm{d}r \, \gamma(r) \frac{\sin qr}{qr} \tag{2.1}$$

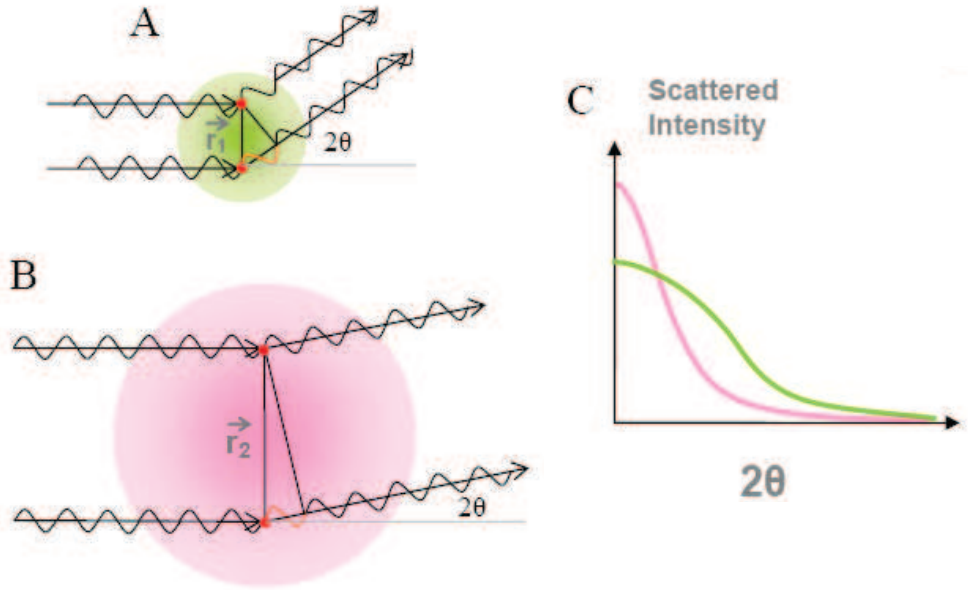

Figure 2.1: This diagram illustrates the addition of X-rays scattered coherently from spheres of different sizes. The path length difference between different parts of the scattering object and the observation point creates interference which will be most intense at the lowest scattering objects. The interference of small dimensions produces more signal at higher angles, as shown. Altogether the result is a scattering pattern which is characteristic of the object. To the right is a comparison of patterns from large and small spheres. These pictures were originally published in Glatter and Kratky [1] and were colorized by Kwok [2].

was worked out by Debye and Bueche [3]. V is the particle volume and $\gamma(r)$ is a correlation function of the intraparticle distances. The scattering angle is usually defined in terms of the momentum transfer vector, $q = 4\pi\frac{\sin(\theta)}{\lambda}$. This definition has the advantage of being independent of the wavelength of the probing X-ray beam.

2.2 Information collected via SAXS

2.2.1 Conventional analyses

While the exact solution of equation 2.1 is complex and dependent on the shape of the molecule under observation, an approximation for the smallest measurement angles was discovered by Guinier [4], for whom it is named. The Guinier equation,

$$I(q) = I(0)\, e^{-\frac{(R_g q)^2}{3}} \tag{2.2}$$

introduces the radius of gyration, R_g, a measure of average particle size[1]. This quantity can be extracted with a linear fit to a plot of $\ln(I(q))$ versus q^2 in the range $R_g q < 1.3$. In practice, unless the data are very noisy, the appropriate region can easily be identified by test fitting several ranges of q. After discovering $q < \frac{1.3}{R_g}$ this way, a more exact R_g can be calculated. Also, if the incident beam intensity is well known, an estimate of molecular weight can be derived from $I(0)$.

Information beyond these quantitative parameters can be gathered from the shape of the scattering curve. General structures such as spheres, cylinders or very extended molecules might be recognized from the characteristic shape of the scattering curve[1]. Subtle differences between related molecules can be elucidated from comparisons of scattering data. However, it is often more intuitive to look at the Fourier transform of the data to real space. The distance distribution function produced is generally referred to as a $P(r)$. A direct transform of the data is impossible because data cannot be acquired at $q = 0$, due to overlap with the probe X-ray beam. Thus it becomes necessary to adopt an indirect

method which is outlined in Glatter and Kratky [1].

The program GNOM[5] provides a numerical implementation of the indirect transform and has come into widespread use. This program employs the regularization method[6] which relies on a stabilizing function to accurately solve the ill-posed problem of transforming data which is not smooth and incomplete. It also requires the input of the maximum molecular dimension, D_{max}, by the user. GNOM users typically find this number by testing several solutions and choosing the one which maximizes the regularization parameters, chosen to keep the solution both true to the data and stable. Often the method works quite well, but between the need to determine D_{max} and the appropriate weight of the stabilization function, each solution must be carefully considered.

2.2.2 Reconstructions

For many years information which could be extracted from scattering data was limited to the two-dimensional patterns and molecular constants mentioned above. Three dimensional interpretation was done by comparison to calculated scattering from known structures or basic geometric shapes. The idea of seeking a molecular shape which could match the scattering curve exactly was mentioned in Glatter and Kratky [1], but the computation required was prohibitive at the time.

Since the late 1990s a variety of computational methods have been used to generate low-resolution molecular shape envelopes from SAXS data with limited *a priori* information. Such programs include SASHA[7], DALAI_GA[8], SAXS3D[9] and SASMODEL[10]. Today both DAMMIN[11] and GASBOR[12], avail-

able from the Biological SAXS group at the European molecular biology laboratory, are widely used. GASBOR is optimized for small proteins and uses a genetic algorithm to "move" a fixed number of "residues" entered by the user to find a model with a shape that matches the data. DAMMIN is more generally applicable and uses a Monte Carlo approach to adjust the solution density of dummy atoms in fixed locations within a sphere of fixed diameter to the same end. Solutions from these methods are not necessarily unique, but if this computation process is repeated a number of times, the set of shapes are often found to agree to within a small deviation.

2.3 Time-resolved SAXS measurements

SAXS measurements can be performed in conjunction with a number of techniques used to make time-resolved measurements. Often such methods are originally developed for use with some sort of optical probe and later adapted for use with SAXS. Because X-rays interact with matter very differently than visible light, a number of adjustments must be made to an apparatus. A few of these techniques are discussed below.

Mixing methods induce molecular action by altering the solvent environment and subsequently probing the sample to monitor changes[13;14]. In the stopped-flow approach, fluids are directed into a turbulent mixing region and then passed into an observation volume. At this point the fluid flow is suddenly stopped and data may be collected at any time after mixing. Time after mixing is established from the sample flow rate, volume between the mixing area and sample cuvette and the delay between stopping the flow and data collection.

Such mixers are commercially available from companies such as KinTek (Texas) and Biologic (France). Use for SAXS requires a specialized sample cuvette and carefully timed control of X-ray exposures. Precise instructions for implementing such a setup at the G1 beamline at CHESS are given in Appendix A.

Continuous flow mixers can be used to induce changes in pH and salt concentration. Both rapid diffusion mixers[15;16] and turbulent T mixers[17;18] have been implemented in SAXS measurements. These have particular advantages for SAXS. The method of data collection with these mixers allows the separation of time after mixing from X-ray exposure time. This allows collection of a fast time-resolved point with a long X-ray exposure which is often necessary to produce sufficient signal-to-noise. It also prevents X-ray damage to the sample because new sample is constantly flowed through the probe beam.

Both the stopped-flow and rapid diffusion techniques are used for some of the measurements described in this document. We have also performed photoexcitation in a continuous flow cell. This technique was developed for use with spectroscopic techniques by Toepke et al. [19]. The modifications needed to use such a device with SAXS are discussed below.

2.3.1 Important parameters for continuous flow measurements

The use of a continuous flow cell in the time-sensitive measurement of a chemical reaction confers the general advantage of separating two relevant and interdependent timescales: time after reaction initiation and length of data collection. This can allow for fast measurements when poor signal to noise of real time data collection would be prohibitive. Measurement time is determined by the aver-

age linear fluid flow speed $\bar{u}$ and the spatial separation between the initiation of the reaction x, resulting in an average time after reaction of

$$\bar{t} = \frac{x}{\bar{u}}. \tag{2.3}$$

This concept is illustrated in Figure 2.2.

Fluid flow is generally categorized as either laminar or turbulent. While the Reynolds number, determined from fluid and flow parameters[20], can be used to roughly predict which type of flow will be present in a system, direct observation is required for conformation. The laminar flow regime is characterized by a continuous fluid velocity profile with the boundary condition velocity equal to zero at the edges of the channel. For flow through a circular pipe, the fluid velocity, $u(r)$ as a function of radial position in the pipe r can be written as

$$u(r) = \frac{G}{4\mu}(a^2 - r^2)\,^{20} \tag{2.4}$$

$$\bar{u} = \frac{Ga^2}{8\mu}\,^{20} \tag{2.5}$$

$$u(r) = 2\bar{u}\left(1 - \frac{r^2}{a^2}\right) \tag{2.6}$$

where a is the diameter of the pipe, G is the pressure gradient of the fluid stream, and μ is the fluid viscosity. The parabolic flow profile described in this equation is illustrated in Figure 2.3.

An X-ray beam used to probe a section of a laminar flow path will sample all the molecules in the intersection of the beam and the pipe in the measurement, including the slow moving ones near the edge of the pipe. Thus the ensemble average will lack precision due to the wide range of sample ages in the volume probed. To minimize this effect it is desirable to flow a sheath of material without sample at the outside of the channel. We can compute the average flow

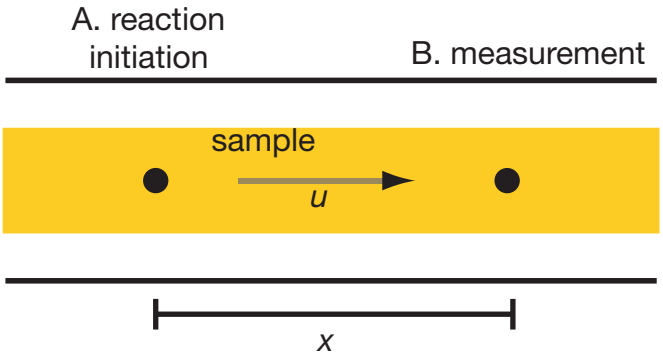

Figure 2.2: At point A, a reaction is initiated. This may be a defined mixing point, a contact point for external stimulus, or the location of some other interaction. The sample is continuously flowing, and is then probed at point B. The distance between the two points, x, is proportional to the time after the reaction.

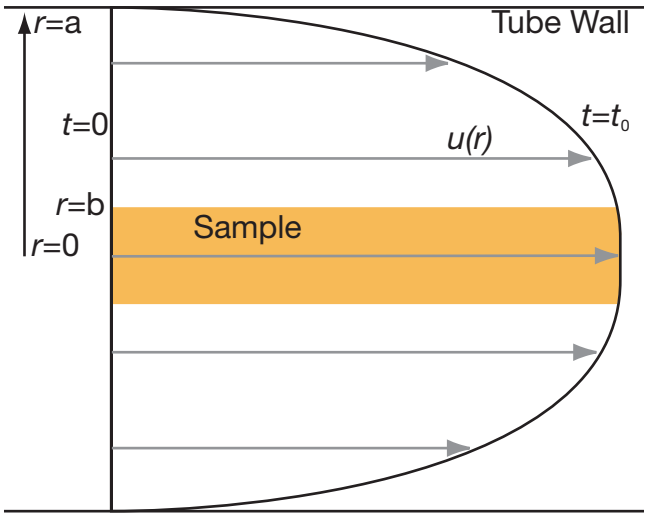

Figure 2.3: Laminar fluid flow has a speed profile which resembles a parabola with the fluid near the channel walls having the boundary condition $u(a) = 0$.

speed of an inner section of the pipe flow with radius b ($\bar{u}_b$) by integrating the fluid speed over the area.

$$\bar{u}_b = \frac{2\bar{u}\pi \int_0^{2\pi} \int_0^b \left(1 - \left(\frac{r}{a}\right)^2\right) r \, dr d\theta}{\pi b^2} \tag{2.7}$$

$$= \bar{u}\left(2 - \left(\frac{b}{a}\right)^2\right) \tag{2.8}$$

Because we control the volume speed of fluids into the flow cell, $\bar{v}_{total}$ and

$\bar{v}_{sample}$, the proportion of the sample flow speed to the total flow speed, k, such that $\bar{v}_{sample} = k\bar{v}_{total}$ is also managed, and $\bar{u}$ can be easily calculated by dividing by the area of the channel. Thus

$$\bar{u}_b = k\bar{u}\left(\frac{a}{b}\right)^2 \text{ and} \tag{2.9}$$

$$\bar{u}\left(2 - \left(\frac{b}{a}\right)^2\right) = k\bar{u}\left(\frac{a}{b}\right)^2 \tag{2.10}$$

and so we can calculate b from known quantities. This leads to the equation

$$b^2 = a^2\left(1 \pm (1-k)^{\frac{1}{2}}\right) \tag{2.11}$$

but because both k and $(1 \pm (1-k))$ are proportionality constants which must be between 1 and 0, $(1-(1-k))$ is the only physical choice. This expression for the radius of the sample can be substituted in equation 2.8, giving

$$\bar{u}_b = \bar{u}\left(1 - (1-k)^{\frac{1}{2}}\right). \tag{2.12}$$

The rate equation can then be used to determine the average time after stimulation. The range of times in the ensemble can be determined from $u(b)$ and $u(0)$.

2.3.2 Considerations for implementing a flow cell in a SAXS setup

There are two particular points in which a flow cell designed for an optical experiment will have to be modified for use in a SAXS setup. The first is mixer size and the second is window material. Both must be changed to account for the different ways in which X-rays pass through matter.

Optimum sample thickness is important in a SAXS experiment where sufficient signal-to-noise is often difficult to achieve. Two factors must be balanced. Thicker samples scatter more. However, they also attenuate the X-ray beam, leading to a loss of intensity. Scattering intensity is proportional to the amount of sample in the path, which will be directly proportional to thickness in these samples. The beam will be attenuated exponentially. Thus the basic equation relating signal intensity to the thickness of the sample is

$$\mathrm{I}(r) = \mathrm{I}_0\, 2r\, e^{-\frac{2r}{\lambda_s}} \tag{2.13}$$

where I_0 is the incident intensity and λ_s is the attenuation length of the sample. The thickness of a tubular cell where the X-rays intersect the center of the cell perpendicular to the tube axis is $2r$, where r is the radius of the tube. If we wish to account for a sheath flow, such that the sample radius is cr, where c is a factor < 1, this becomes

$$\mathrm{I}(r) = \mathrm{I}_0\, 2cr\, e^{-\frac{2cr}{\lambda_s}}\, e^{-\frac{2r(1-c)}{\lambda_w}} \tag{2.14}$$

where λ_w is the attenuation length of the sheath buffer. To find the sample cell radius which maximizes the output intensity, set the derivative of I(r) to zero.

$$0 = 1 - 2r\left(\frac{1}{\lambda_w} + c(\frac{1}{\lambda_s} - \frac{1}{\lambda_w})\right) \tag{2.15}$$

$$2r = \frac{1}{\frac{1}{\lambda_w} + c(\frac{1}{\lambda_s} - \frac{1}{\lambda_w})} \tag{2.16}$$

If the sample is fairly dilute and the sheath solution has otherwise the same composition, the second term will cancel out, and the diameter of the sample tube should optimally be the attenuation length of the sheath solution. Since most buffer solutions are dilute aqueous solutions, the attenuation length of water is an appropriate estimate for the ideal sample cell thicknesses at different energies. Figure 2.4 shows the attenuation lengths of water over a range of energies

commonly used at the G1 station at CHESS[21]. To reduce sample consumption, we chose to develop a 1 mm thick flow cell, which is ideal for an X-ray energy of 8 keV.

Window material is a consideration in any X-ray experiment. Because X-rays interact directly with electrons any sample container will add to the background signal. To minimize this effect, window materials which can be made very thin are often chosen. To enable photoexcitation it is also necessary that the window be optically transparent. Thin-walled polyester tubing is available with very thin, optically transparent walls from Advanced Polymers, Inc. (Vermont)[23]. A quick test with a 355 nm laser and a power meter demonstrated the tubing is also UV transparent at these thicknesses, making it useful in a wide variety of experiments involving photoexcitation.

2.3.3 Flow cell with sheath flow

The flow cell shown in Figure 2.5 was built for the purpose of making time-resolved measurements on optically excitable samples. The sample is brought into the sample cell area through a narrow piece of hypodermic tubing. This allows buffer to flow around the sample, creating a sheath flow. Because the flow cell is made from cylindrical tubing, the excitation and probe beams can enter the cell at different angles, which is essential given the space limitations of the beamline. Machine drawings and precise assembly instructions can be found in Appendix B.

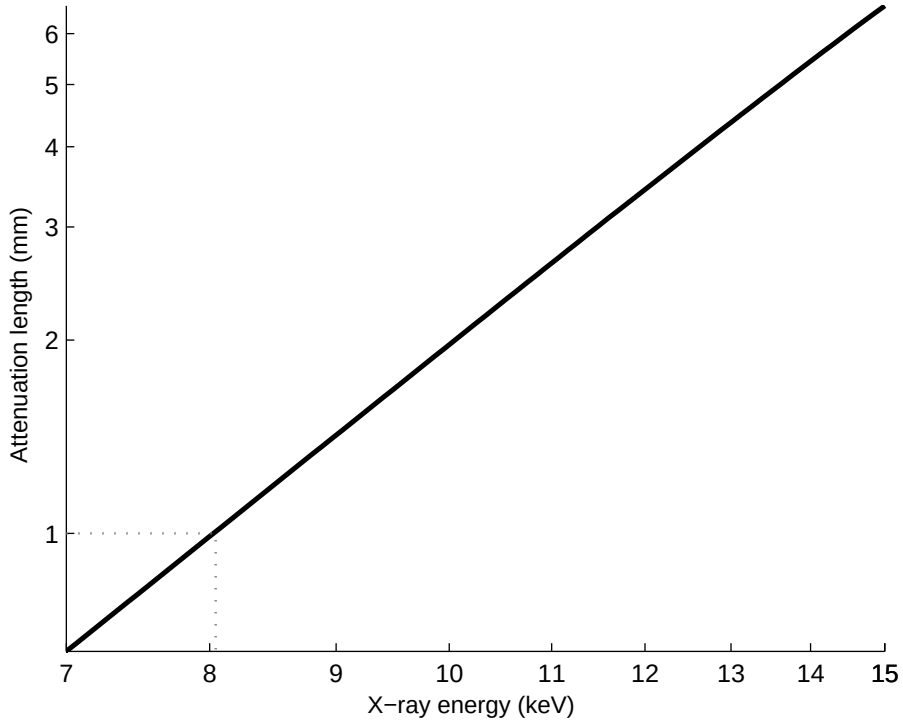

Figure 2.4: Attenuation lengths of water over the energies typically found at the G1 station. The Center for X-ray Optics (CXRO) website calculator was used to generate the data[22].

2.3.4 Setup at G1 station

Figure 2.6 show the SAXS setup at the G1 station of the Cornell High Energy Synchrotron Source (CHESS) which incorporates a laser and a flow cell in a SAXS experiment. Upstream (closer to the X-ray source) of the sample, are three sets of beam defining slits and two ion chambers which are typical of any SAXS setup. A focusing capillary has been employed to create a small beam of sufficient intensity to collect sufficient signal-to-noise. The sample cell is located downstream of the capillary. Mounted on the sample cell is another set of slits to act as guard slits from the capillary aperture. The excitation laser is directed toward the flow cell from above by mirrors, which are mounted along with the laser head on a large stepper motor stage. The remaining sample-to-

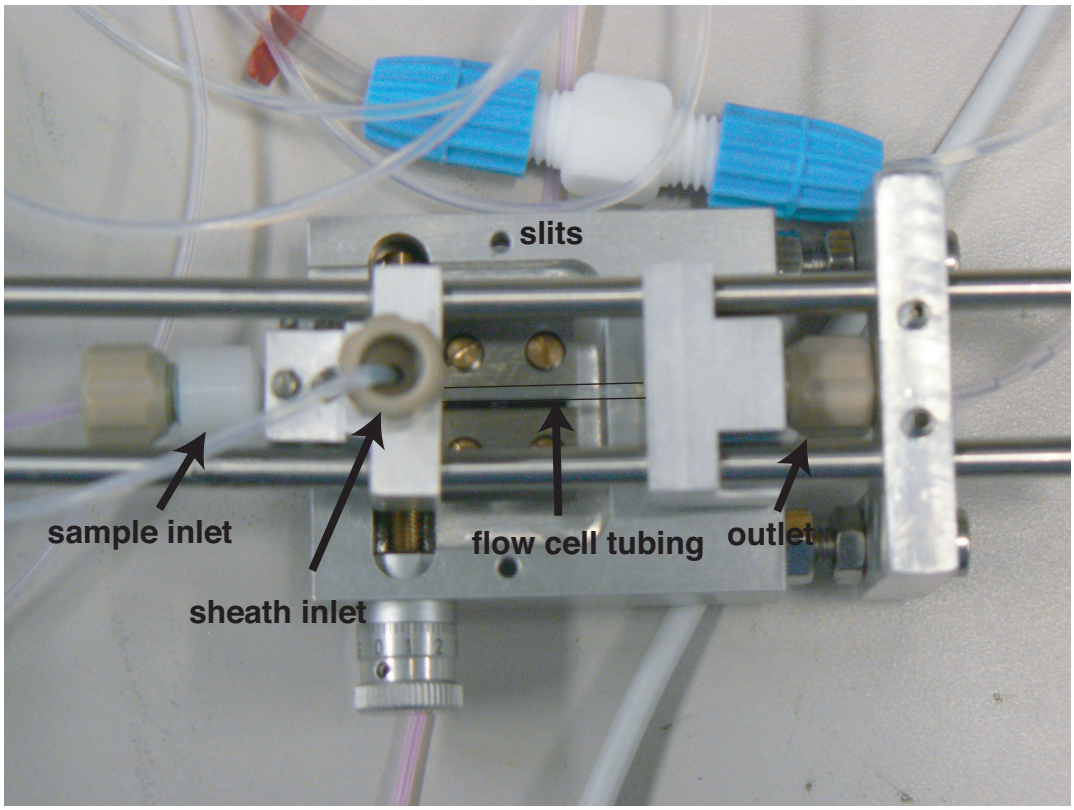

Figure 2.5: Picture of the mounted flow cell. The polyester tubing is outlined because it is very difficult to discern in the photograph.

detector distance is under vacuum, to avoid signal attenuation and increased background from air.

In order to align the sample beam and the capillary, a 20 μm slit is used. This slit is mounted at roughly 45° angle to both the laser and X-ray beams. It is aligned with the X-ray beam, and then the laser beam is steered through it, using pin diodes to detect transmission of both beams through the slit. After this procedure both beams will be coincident in the x direction (as defined in section 2.3.1) and the laser motor position is set to zero. The slit is then replaced with the loaded sample cell. The delay time between excitation and probing with the X-ray beam can be changed by moving the laser beam in x.

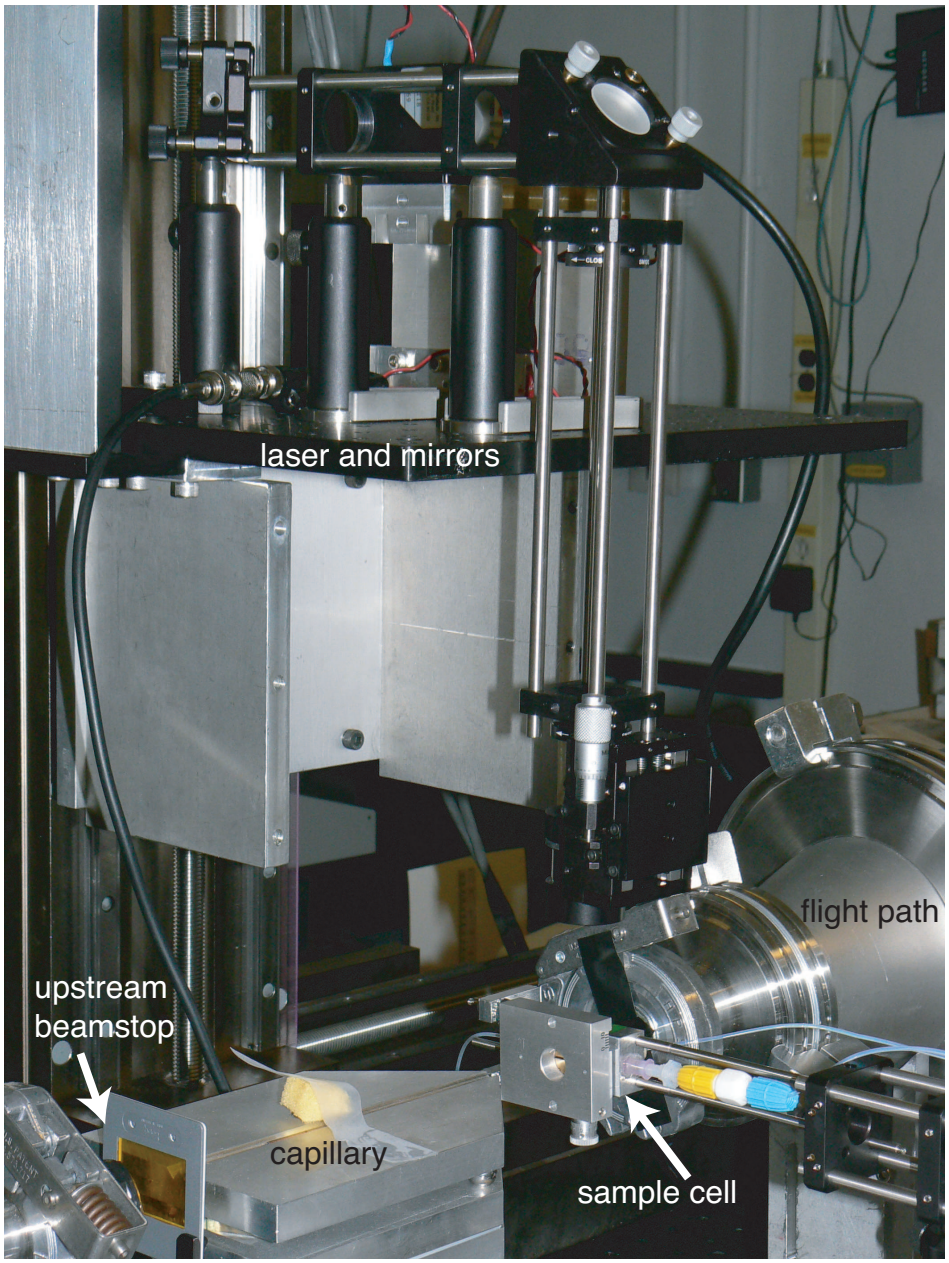

Figure 2.6: Photograph of the setup at G1 station. The picture shows (from left to right) the end of the chamber which houses the defining slits (under He), the upstream beamstop for the focusing capillary, an extra pair of slits on the sample cell, the sample cell with the laser beam coming in from above, and the downstream flight path which is under rough vacuum. In the background is the large motor used to hold and position the laser.

REFERENCES

1. Glatter, O. and O. Kratky, editors. *Small Angle X-ray Scattering*. Academic Press Inc., London, 1982.

2. Kwok, L. *Early Compaction Kinetics of the* Tetrahymena *Ribozyme Measured with Time Resolved Small Angle X-Ray Scattering*. Ph.D. thesis, Cornell University, Ithaca, NY, 2006.

3. Debye, P. and A. M. Bueche. "Scattering by an inhomogeneous solid." *J. Appl. Phys.*, **20**(6): 518–525, 1949.

4. Guinier, A. "La diffraction des rayons X aux tres petits angles; application à l'etude de phenomenes ultramicroscopique." *Ann. Phys. (Paris)*, **12**: 166–237, 1939.

5. Svergun, D. I. "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." *J. Appl. Cryst.*, **25**: 495–503, 1992.

6. Tikhonov, A. N. and V. Y. Arsenin. *Solution of Ill-Posed Problems*. Wiley, New York, 1977.

7. Svergun, D. I., V. V. Volkov, M. B. Kozin, H. B. Stuhrmann, C. Barberato, and M. H. Koch. "Shape determination from solution scattering of biopolymers." *J. Appl. Cryst.*, **30**: 798–802, 1997.

8. Chacón, P., F. Morán, J. Díaz, E. Pantos, and J. Andreu. "Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm." *Biophys. J.*, **74**(6): 2760–2775, 1998.

9. Walther, D., F. E. Cohen, and S. Doniach. "Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules." *J. Appl. Cryst.*, **33**: 350–363, 2000.

10. Vigil, D., S. Gallagher, J. Trewhella, and A. Garcia. "Functional dynamics of the hydrophobic cleft in the N-domain of calmodulin." *Biophys. J.*, **80**(5): 2082–2092, 2001.

11. Svergun, D. "Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing." *Biophys. J.*, **76**(6): 2879–2886, 1999.

12. Svergun, D. I., M. V. Petoukhov, and M. H. J. Koch. "Determination of domain structure of proteins from X-ray solution scattering." *Biophys. J.*, **80**(6): 2946–2953, 2001.

13. Gibson, Q. and L. Milnes. "Apparatus for rapid and sensitive spectrophotometry." *Biochem. J.*, **91**(1): 161–171, 1964.

14. Moody, M. F., P. Vachette, A. M. Foote, A. Tardieu, M. H. Koch, and J. Bordas. "Stopped-flow x-ray scattering: the dissociation of aspartate transcarbamylase." *Proc. Natl. Acad. Sci. U. S. A.*, **77**(7): 4040–3, 1980.

15. Knight, J. B., A. Vishwanath, J. P. Brody, and R. H. Austin. "Hydrodynamic focusing on a silicon chip: Mixing nanoliters in microseconds." *Phys. Rev. Lett.*, **80**: 3863–3866, 1998.

16. Pollack, L., M. W. Tate, N. C. Darnton, J. B. Knight, S. M. Gruner, W. A. Eaton, and R. A. Austin. "Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **96**: 10115–10117, 1999.

17. Takahashi, S., S. Yeh, T. Das, C. Chan, D. Gottfried, and D. Rousseau. "Folding of cytochrome *c* initiated by submillisecond mixing." *Nat. Struct. Biol.*, **4**(1): 44–50, 1997.

18. Akiyama, S., S. Takahashi, T. Kimura, K. Ishimori, I. Morishima, Y. Nishikawa, and T. Fujisawa. "Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **99**: 1329–1334, 2002.

19. Toepke, M. W., S. H. Brewer, D. M. Vu, K. D. Rector, J. E. Morgan, R. B. Gennis, P. J. A. Kenis, and R. B. Dyer. "Microfluidic flow-flash: Method for investigating protein dynamics." *Anal. Chem.*, **79**(1): 122–128, 2007.

20. Tritton, D. J. *Physical fluid dynamics*. Clarendon Press, Oxford, 2nd edition, 1988.

21. 2008. URL `http://www.chess.cornell.edu/gline/index.htm`.

22. 2008. URL `http://henke.lbl.gov/optical_constants/atten2.html`.

23. Kalinin, Y., J. Kmetko, A. Bartnik, A. Stewart, R. Gillilan, E. Lobkovsky, and R. Thorne. "A new sample mounting technique for room-temperature macromolecular crystallography." *J. Appl. Cryst.*, **38**(2): 333–339, 2005.

CHAPTER 3

FOCUSING CAPILLARY OPTICS FOR USE IN SAXS

When using a continuous flow cell with SAXS a small, intense x-ray beam is necessary. A large beam will increase the range of time after excitation in the sample under investigation due to the proportional relationship between x and t, as defined in Chapter 2, but a small beam could also lead to intolerably poor signal-to-noise. This chapter describes a means of answering both issues. It is reproduced with minimal alteration from a laboratory note previously published in *The Journal of Applied Crystallography*[1].

3.1 Introduction

Borosilicate glass capillaries have been used to focus x-ray beams into high flux spots as small as micrometers across[2] at the expense of beam collimation. Recently, continued development and widespread use of these focusing capillaries has been pursued at the Cornell High Energy Synchrotron Source (CHESS)[3;4] in applications such as Laue diffraction, x-ray imaging[3], and crystallography[5]. Capillaries have also been tested in scattering studies involving wide and small angle x-ray scattering[6;7].

SAXS measures the low angles of the x-ray scattering profile to characterize the size and shape of a collection of randomly oriented molecules, and can be adapted for a wide variety of targets. A focused microbeam has proven useful on small and spatially resolved samples[6;8] and microfluidic devices[9]. Macromolecular folding can be induced using stopped-flow or continuous flow mixers, allowing time-resolved compaction studies[10;11]. In the latter, time is a func-

tion of position, so a large beam may limit time resolution. If the beam flux is low, long exposure times are necessary, requiring increased sample consumption. The high flux microbeam produced by a single bounce focusing capillary is an ideal tool for these experiments, provided the divergence of the beam does not limit the measurement, either by smearing the SAXS profile or restricting access to the small angles needed for the measurement.

We present here SAXS measurements of the protein cytochrome *c* taken with focusing capillaries at the CHESS D1 station. Scattering profiles from a calibrant and cytochrome *c* demonstrate no great loss of data or resolution when compared to profiles acquired with unfocused beam. By using single-bounce capillaries and limiting the accepted beam, a microbeam with sufficiently low divergence to allow SAXS on moderately sized proteins can be generated. These capillaries are readily incorporated into existing SAXS setups.

3.2 Implementation

3.2.1 Sample

Cytochrome *c* was obtained from Sigma and used without further purification. Cytochrome *c* was dissolved at 30 mg/mL in a pH 7 citrate-phosphate buffer. Sample was pumped at a rate of 0.1 mL/min through a small tube consisting of a 2 mm diameter thin-walled polyester (PET) tube (Advanced Polymers, Salem, NH)[12]. Silver stearate (d spacing 48.68 Å[13]) powder held in a 1 mm PET tube was used as a calibrant.

3.2.2 Focusing Capillary

The single bounce focusing capillary identified as BSG644 was used for these measurements. The inner diameters of the base and the tip of the capillary are 400 μm and 266 μm respectively, and the length is 11 mm. Using the D1 beam source, we expect the capillary to produce a 15 μm spot[14] located 52 mm from the tip. The angular divergence with the optic fully illuminated is 4.3 mrad; guard slits restricted the beam incident on the capillary (as shown in Figure 3.1) to limit the divergence to approximately 2 mrad. To estimate the size of a molecule which might scatter below 2 mrad, we note the D spacing of a lattice that would scatter to 2 mrad is given by $D = \frac{\text{wavelength}}{\text{angle}}$[15]. At 9 keV spatial variations on length scales larger than 690 Å cannot be probed. Fully unfolded cytochrome *c* has a radius of gyration (R_g) of only 32 Å[16], which is much smaller than the limiting value. Thus, we do not expect the divergence of the beam to be an issue for measurements on small proteins.

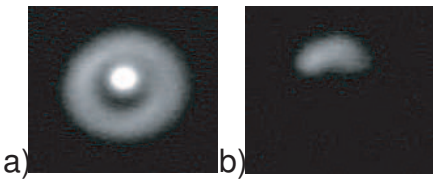

Figure 3.1: Images of the beam incident on a fluorescent screen approximately 1 m from the focus. Image a shows the entire beam, including the unfocused beam through the capillary. Image b shows the portion of the beam used in the measurements. The unfocused beam has been blocked, as well as all but one quadrant of the focused beam to minimize divergence.

3.2.3 Beamline

Measurements were conducted at the D1 bending magnet station. A beam energy of 9 keV with 1.5 % resolution was achieved through double bounce multilayers. Figure 3.2 depicts the station setup, typical for SAXS, with the exception of the focusing capillary and an additional 1 mm pinhole. This guard aperture was attached to the sample cell, and eliminated most parasitic scatter from the capillary tip as well as the scatter from the air around the capillary. A small amount of parasitic scatter and small angle scatter from the glass capillary tip had to be masked in the analysis process. The sample was placed at the focal spot of the capillary.

The capillary was mounted in a groove on a motorized stage which controlled the position and the tilt angle. These were carefully adjusted until a far field image produced on a fluorescent screen showed a centered beam. The upstream slits were then set to block the unfocused beam and further restrict the divergence of the focused beam (Figure 3.1). The incident beam was positioned on the lower part of the capillary.

To explore the effect of focusing on the SAXS profiles, we acquired data without the capillary. For these measurements, the beam was defined to be 0.1 mm $\times$ 0.1 mm with slits and the capillary was moved out of the beam path. The setup was not otherwise altered.

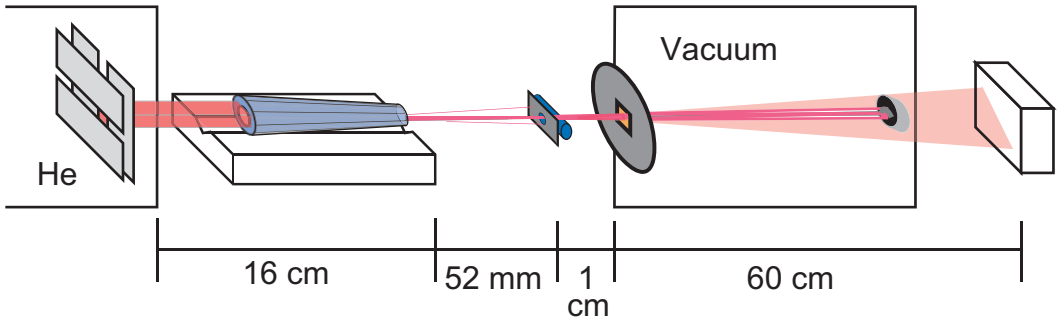

Figure 3.2: Downstream beamline setup (not to scale): depicted are the slits (in He) the focusing capillary, pinhole, and sample (in air), the downstream flight path with a pin diode incorporated into the beamstop, and finally the CCD.

3.3 Results

A pin diode monitor was placed in the beam stop to monitor the x-ray flux, both with and without the capillary. We measured 30% higher flux with the capillary in place. However, the slit-defined beam had an area of 100μm $\times$ 100μm while the calculated focal spot for the capillary is much smaller, about 180 μm^2. Thus beam from the capillary is about 60 times more intense.

This intensity gain is easily seen by comparing 10 second exposures of silver stearate, taken with and without the capillary (Figure 3.3). The relative heights of the calibrant peak at q = 0.129 Å^{-1} (where $q = \frac{4\pi \sin\theta}{\lambda}$) reflects the gain in intensity. Furthermore, the full width half max of the peaks goes from .0053 Å^{-1} with the capillary to .0046 Å^{-1} without it, showing only a small amount of smearing.

The scattering profile of cytochrome *c* was also measured with and without the capillary. Four 60 second exposures were acquired under each condition,

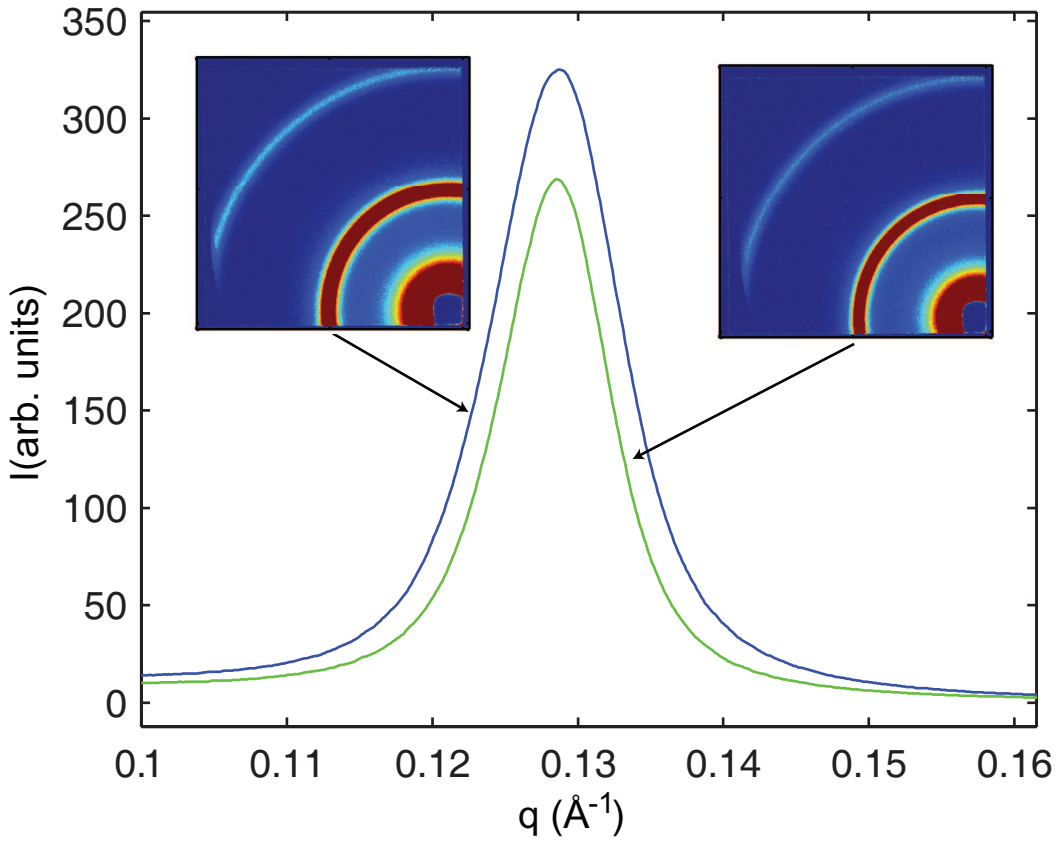

Figure 3.3: Comparison of the silver stearate image taken with a capillary focused beam, and with the beam defined by slits. The main plot is the radially integrated SAXS profile of the first ring for both measurements; the insets show the CCD images. The line in blue (upper curve) represents the image taken with the focusing capillary, the one in green with slits. The peaks are 324 and 267 intensity units respectively.

two with protein, and two without for subtracting the scattering background. When the curves are scaled to account for the higher intensity of the capillary beam, the curve shapes are seen to agree within the noise present in the data. Thus, information about the size and shape of the protein in solution are not altered by the focusing optic.

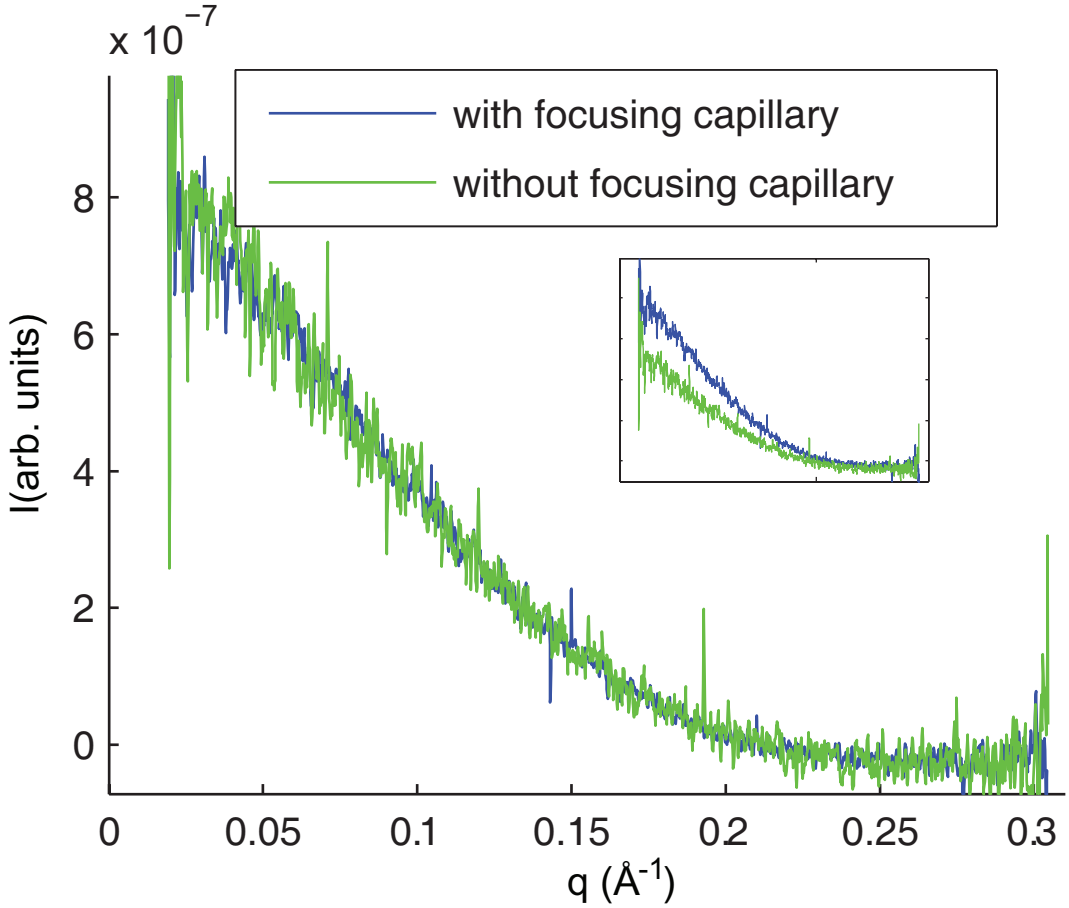

Figure 3.4: Radially integrated SAXS intensity profiles of cytochrome *c*; the curve taken without the capillary has been adjusted to account for the difference and beam intensity and a slight detector offset. The profiles without adjustments are shown in the inset. The lowest q value reflects the size of the beamstop, the highest value the end of the detector range.

3.4 Conclusion

We have demonstrated that focusing capillaries with mrad of divergence can be useful in solution SAXS studies on small molecules. Many biological molecules are small enough so that the low divergence from these capillaries will not block important scattering lengths. The use of a small, highly intense beam enables acquisition of good quality scattering data from small sample volumes. In conjunction with continuous flow mixers, sharper time resolution can be achieved

compared to slit-defined beams. In the future we intend to use this setup to take time resolved data.

We would like to thank Ernie Fontes at CHESS for his assistance with these measurements. This work was supported by NASA, NIH and NSF. This work is based upon research conducted at the Cornell High Energy Synchrotron Source (CHESS) which is supported by the National Science Foundation and the National Institutes of Health/National Institute of General Medical Sciences under award DMR-0225180.

REFERENCES

1. Lamb, J. S., S. Cornaby, K. Andresen, L. Kwok, H. Park, X. Qiu, D. M. Smilgies, D. H. Bilderback, and L. Pollack. "Focusing capillary optics for use in solution small-angle X-ray scattering." *J. Appl. Cryst.*, **40**: 193–195, 2007.

2. Engstrom, P., S. Larsson, A. Rindby, A. Buttkewitz, S. Garbe, G. Gaul, A. Knochel, and F. Lechtenberg. "A submicron synchrotron x-ray beam generated by capillary optics." *Nucl. Instrum. Methods Phys. Res., Sect. A*, **302**: 547–552, 1991.

3. Bilderback, D. H., S. A. Hoffman, and D. J. Thiel. "Nanometer spatial resolution achieved in hard x-ray imaging and Laue diffraction experiments." *Science*, **263**: 201–203, 1994.

4. Huang, R. and D. H. Bilderback. "Single-bounce monocapillaries for focusing synchrotron radiation: modeling, measurements and theoretical limits." *J. Synchrotron Rad.*, **13**: 74–84, 2006.

5. Bilderback, D. H. and R. Huang. "X-ray tests of microfocusing monocapillary optic for protein crystallography." *Nucl. Instrum. Methods Phys. Res., Sect. A*, **467-468**: 970–973, 2001.

6. Riekel, C. "Applications of micro-SAXS/WAXS to study polymer fibers." *Nucl Instrum Meth B*, **199**: 106–111, 2003.

7. Riekel, C., M. Burghammer, and M. Müller. "Microbema small-angle scattering experiments and their combination with microdiffraction." *J. Appl. Cryst.*, **33**: 421–423, 2000.

8. Zafeiropoulos, N. E., R. J. Davie, S. V. Roth, M. Burghammer, K. Schneider, C. Riekel, and M. Stamm. "Microfocus X-ray scattering scanning microscopy for polymer applications." *Macromol. Rapid. Comm.*, **26**: 1547–1551, 2005.

9. Otten, A., S. Koster, B. Struth, A. Snigirev, and T. Pfohl. "Microfluidics of soft matter investigated by small-angle X-ray scattering." *J. Synchrotron Rad.*, **12**: 745–750, 2005.

10. Pollack, L., M. W. Tate, N. C. Darnton, J. B. Knight, S. M. Gruner, W. A. Eaton, and R. A. Austin. "Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **96**: 10115–10117, 1999.

11. Semisotnov, G. V., H. Kihara, N. V. Kotova, K. Kimura, Y. Amemiya, K. Wakabayashi, I. N. Serdyuk, A. A. Timchenko, K. Chiba, and K. Nikaido. "Protein globularization during folding. A study by synchrotron small-angle X-ray scattering." *J. Mol. Biol.*, **262**: 559–574, 1996.

12. Kalinin, Y., J. Kmetko, A. Bartnik, A. Stewart, R. Gillilan, E. Lobkovsky, and R. Thorne. "A new sample mounting technique for room-temperature macromolecular crystallography." *J. Appl. Cryst.*, **38**(2): 333–339, 2005.

13. Vand, V., A. Aitken, and R. K. Campbell. "Crystal structure of silver salts of fatty acids." *Acta Cryst.*, **2**: 398–403, 1949.

14. Huang, R. "Single-bounce capillary focusing calculator." 2005. URL `http://glasscalc.chess.cornell.edu/imageprof.html`.

15. Guinier, A. and G. Fournet. *Small-Angle Scattering of X-Rays*. John Wiley and Sons., Inc., New York, 1955.

16. Kataoka, M., Y. Hagihara, K. Mihara, and Y. Goto. "Molten globule of cytochrome *c* studied by small angle X-ray scattering." *J. Mol. Biol.*, **229**: 591–596, 1993.

CHAPTER 4

RECONSTRUCTING THREE DIMENSIONAL SHAPE ENVELOPES FROM TIME RESOLVED SMALL ANGLE X-RAY SCATTERING DATA

To extract the most information from a SAXS experiment, we make use of a wide assortment of analysis techniques. Because reconstructions are becoming a widespread tool for interpretation of static SAXS data, we wished to asses their utility in the time-resolved experiments previously mentioned. To that end we carried out a number of tests. This chapter is taken with minimal alteration from a paper that is in press at the *Journal of Applied Crystallography*[1].

4.1 Introduction

Solution small angle x-ray scattering (SAXS) provides low-resolution structural information about biomolecules. Geometrical parameters such as radius of gyration (R_g), fractal dimension, or surface to volume ratio, are extracted from SAXS data by examining different regions of the scattering curve[2]. Additional structural information is inferred through comparison of the scattering curve, I(q), or its Fourier inverse, p(r), with models and other data. However, in recent years modern computing power has made it possible to reconstruct a three dimensional molecular shape whose scattering coincides with measured profiles over the entire range of a scattering curve[3–6]. In addition, these *ab initio* algorithms allow users to propose molecular shapes without relying on previous knowledge. These reconstruction programs have produced shape envelopes for many proteins[7;8] and nucleic acids[9;10], enabling straightforward comparison between SAXS measurements and structural models.

Conformational changes in biomolecules, including folding, can be induced by solvent exchange through mixing. Commercially available stopped-flow mixers as well as continuous flow microfluidic mixers have been interfaced with SAXS instrumentation to acquire time resolved measurements of events such as protein and RNA folding[11–16]. Such data are traditionally presented as sets of curves, showing scattering intensity as a function of scattering angle, acquired at different times after the initiation of folding. Alternately, the time dependence of a single parameter, such as R_g, may be used to represent the results.

The information provided by time resolved SAXS measurements, such as the global structures of transient conformational states, is not readily accessible to other techniques. The goal of this study is to evaluate the limits of applicability of reconstruction methods to time resolved data by providing examples and discussing the pros and cons of this approach, in particular addressing two primary concerns. First, due to practical limits on sample consumption, time resolved data are noisy. Noise can mask structural information or worse, be interpreted as actual curvature leading to inaccurate features in the reconstruction. Second, SAXS is an ensemble measurement, naturally sampling all molecular states present in solution. It is unclear how a single reconstruction reflects the mixture of states present in a solution, for example when folded and unfolded molecules coexist.

In an attempt to answer these questions, we reconstructed time resolved data for two related systems. The P4-P6 subdomain of the group I intron from *Tetrahymena Thermophila* is a relatively simple RNA. Though studies of folding which probe local structure reveal additional subtle features, the large scale process reported by SAXS occurs as a single, two state collapse[17]. The entire

Tetrahymena Thermophila ribozyme has a more complex fold[18]. Both molecules have been extensively studied[19;20], and their folded states crystallized in full and in large part, respectively[21;22], making them ideal examples of simple and complex molecules. For both constructs we have collected time-resolved scattering data and reconstructed structures along the Mg^{2+} induced folding pathways. More traditional analyses of these data have provided insight into folding trajectories and have been previously published[17;23]. Here we present reconstructions of these data and compare them with shapes derived from crystal structures of the molecules, to assay the value of reconstruction methods in conjunction with the changing molecular shapes that accompany folding.

4.2 Methods

4.2.1 Sample Preparation

The L-21 *sca* I construct of the *Tetrahymena* ribozyme and P4-P6 subdomain were prepared by *in vitro* transcription as described in Russell and Herschlag [24]. Prior to mixing with $MgCl_2$, full length constructs were stored in a solution of 10 mM K^+MOPS pH 7.0 plus 100 mM KCl and P4-P6 samples were stored in 50 mM K^+ MOPS buffer pH 7.0. All RNA samples were annealed shortly before data collection by heating to 90° C for one minute. This procedure ensured an initial state absent of tertiary structure.

4.2.2 X-ray data collection

APS

Data for both constructs were acquired on time scales ranging from 1.3 to 168 ms at the 8-ID I beamline at the Advanced Photon Source (APS) using a microfluidic mixer as described in Kwok et al. [23]. Tertiary structure acquisition was initiated by diffusive mixing with a solution containing 10 mM $MgCl_2$ in addition to buffering ions matching those in the RNA solution. Up to twenty-five thirty second images of RNA scattering were acquired at each time point, along with a comparable number of buffer background images. The large number was required to ensure sufficient signal-to-noise given the micrometer scale sample size of the jet containing RNA.

CHESS

Time points 157 ms and longer were collected at the G1 station of the Cornell High Energy Synchrotron Source (CHESS) using a modified SFM-4 stopped-flow mixer from Biologic (France). The details are also described in Kwok et al. [23]. Four 100 μs images were acquired for each sample, along with buffer background images. Due to aggregation of P4-P6 on the 100 ms time scale[17], only the full length construct was studied by stopped-flow.

4.2.3 Data Analysis

Initial Data Reduction

Initial data reduction was carried out using Matlab following procedures described in Kwok et al. [23]. Images were radially integrated to yield scattering intensity versus the momentum transfer vector $q = 4\pi\frac{\sin\theta}{\lambda}$, where 2θ is the scattering angle. Multiple scattering curves acquired at each time point, as outlined above, were checked for consistency before averaging and subtraction of the solution background. Portions of the curves deemed unreliable at the lowest q due to parasitic scatter and at the highest q due to low signal to background were removed before beginning reconstructions. The radius of gyration (R_g) was calculated for each curve from the scattering at the lowest angles[2].

Reconstructions

Reconstructions were produced using software made available by the Biological Small Angle Scattering group at the European Molecular Biology Laboratory[25]. GNOM[26] generates a smooth representation of scattering intensity as a function of q, which is required as input for the reconstruction program. In this process the user must specify a maximum dimension (D_{max}). We generally set this parameter to be 5 Å larger than the number which maximized the default regularization parameters which GNOM uses to quantify a reasonable profile. This practice ensures the shape is reconstructed within a sufficiently large sphere. The reconstruction program DAMMIN[3] was run 10 times in slow mode on the Cornell Center for Materials Science 30 node cluster. The runs, which required 12 to 24 hours, were processed in parallel so the entire opera-

tion required only one day. The three dimensional results were averaged using DAMAVER. In addition to an averaged reconstruction, DAMAVER produced a quantitative measure of the agreement between the ten individual shapes, the mean normalized spatial discrepancy (MNSD). A small MNSD indicates consistency between the ten individual reconstructions, and thus a unique solution to the scattering curve. Finally, the averaged bead models were processed by Situs[27] to generate a single surface envelope.

4.3 Results

The P4-P6 domain folds in isolation, independent of the remaining structural elements of the *Tetrahymena* ribozyme. Cate et al. [21] reported the crystal structure and Lipfert et al. [28] presented static reconstructions of this $\approx$ 160 nucleotide domain. Time resolved SAXS studies have shown that global folding is accurately described by a two state model, in which linear combinations of the folded and unfolded scattering curves can accurately reproduce the scattering of the intermediate time points[17]. A plot of R_g versus time after mixing with Mg^{2+} fits a single exponential equation, consistent with a single phase collapse of the molecule (Figure 4.1). Crystal structures reveal that the major feature of the folded state is a 150° bend near the middle of the RNA helix. Electrostatic repulsion in the initial, low salt state restricts the shape of the RNA molecule to extended conformations[18]. Thus P4-P6 folding as observed by SAXS can be treated as a two-state process where, upon the addition of Mg^{2+}, P4-P6 folds nearly in half. This bent structure is stabilized by native tertiary contacts[17].

We reconstructed shapes reported by each time resolved SAXS profile ac-

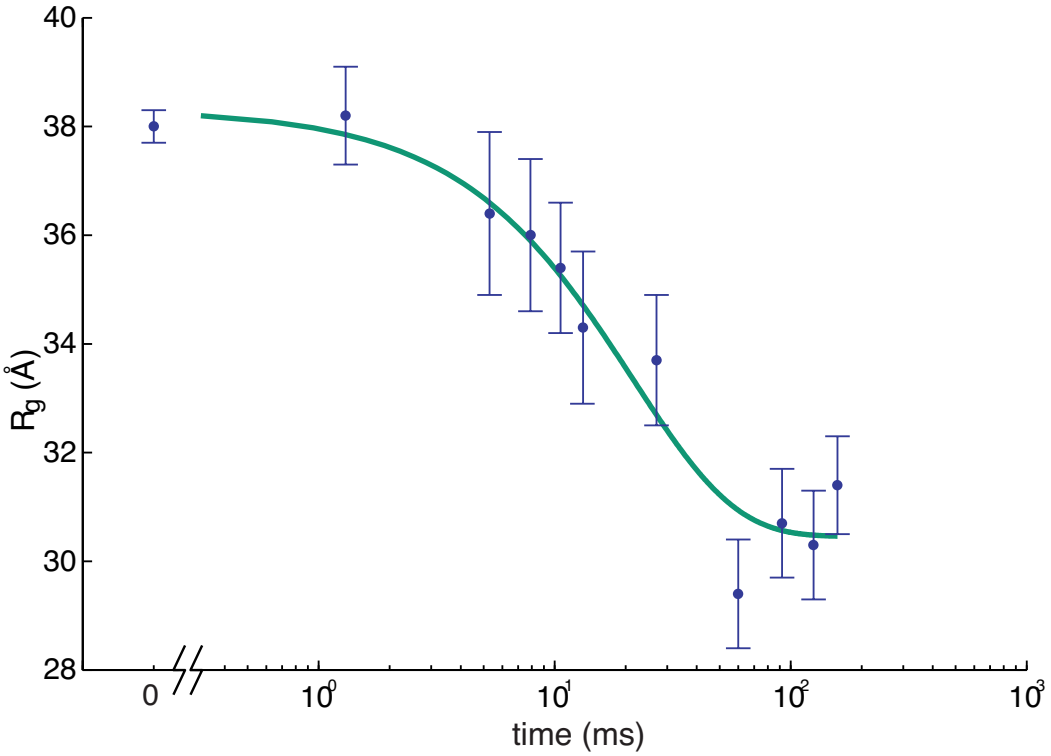

Figure 4.1: Radius of gyration (R_g) versus time for the P4-P6 subdomain. The points are calculated from fits to the Guinier region of the data[2], with the errors being determined by the 95 % confidence intervals. The black line is the fit of the equation $R_g = (L_0 - L_1)e^{\frac{-t}{\tau}} + L_1$ with $L_0 = 38$, $L_1 = 30.5$, and $\tau = 21$ ms. When measured by SAXS, P4-P6 folding is accurately described by a single exponential, thus the global compaction of P4-P6 is apparently a two state process.

quired during P4-P6 folding. The initial SAXS data for these reconstructions are presented in the supplementary material of Schlatterer et al. [17]. These structures, pictured in Figure 4.2 indicate a progression from a thin, elongated structure to a more compact state, consistent with the folding model described above. Additionally, the final state can be docked into the crystal structure[21] as shown in Figure 4.3. Agreement between the ten individual reconstructions for each curve, indicated quantitatively by the MNSD, is good. The large scale features of the reconstructions reproduce well over the set; however, smaller features vary from one reconstruction to the next. A typical comparison of the theoreti-

cal scattering curves of the reconstructions to the GNOM fit and the data, shown in Figure 4.4, demonstrates the fit of both GNOM and the reconstructions to the data. To quantify the agreement we computed the signal-to-noise ratio (S/N) as

$$S/N = \frac{\sum_i fit_i}{\sum_i |fit_i - data_i|}$$

where the output from GNOM was used as the fit. While the overall S/N was 27, quite good, the same analysis done on subsections of the data makes clear the lower S/N in the high q range, as demonstrated in Table 4.1. The noise in the data manifests itself as minor features in the curves output by GNOM. These features are unlikely to represent physical structure, but DAMMIN fits every feature in the GNOM curve precisely (Figure 4.4). To compare calculated scatter from DAMMIN to the GNOM curve and the data, we define the mean discrepancy (MD) between two curves, as

$$MD = \frac{1}{N} \sum_{i=1}^{N} |f_i - g_i|.$$

where f and g are two curves and N the total number of data points in each curve. The MD between the data and the GNOM fit is 3.2×10^{-6}, but the MD between the GNOM fit and the DAMMIN reconstruction is 8.6×10^{-8}, over an order of magnitude smaller. This difference demonstrates the strong dependence of the reconstructions on the GNOM interpretation of the data.

Of particular note is the reconstruction of the scattering curve taken 27 ms after the initiation of folding. This data point was acquired near the midpoint of the compaction illustrated in Fig 4.1, and according to a two state fit, is approximately 63 % unfolded[17]. The structure is not as elongated as that of unfolded P4-P6 but longer and thinner than the fully folded construct. A single half folded P4-P6 molecule would probably have the ends of the helix sepa-

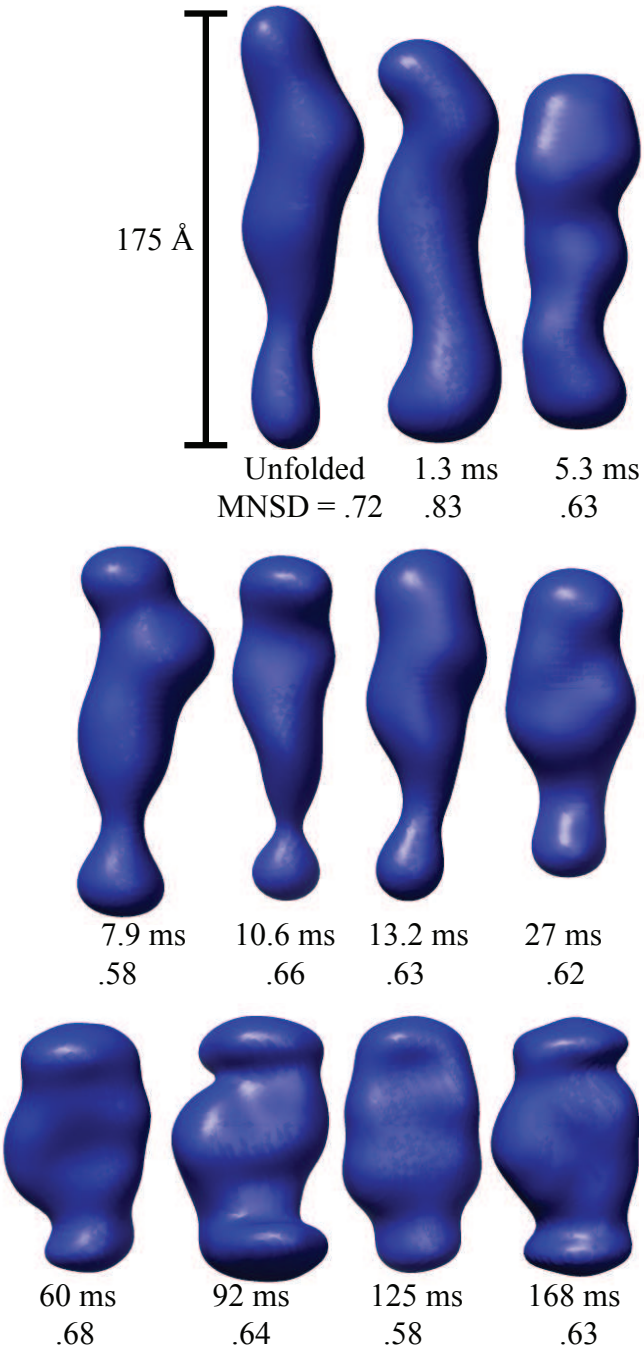

Figure 4.2: Averaged reconstructions of the P4-P6 subdomain. Each shape is labeled with the time elapsed since mixing and the MNSD value that reflects the uniqueness of the reconstruction. While the MNSDs of two are larger than 0.7, examination of the individual reconstructions (not shown) confirms that the averages represent the ensembles. All reconstructions are depicted on the same spatial scale.

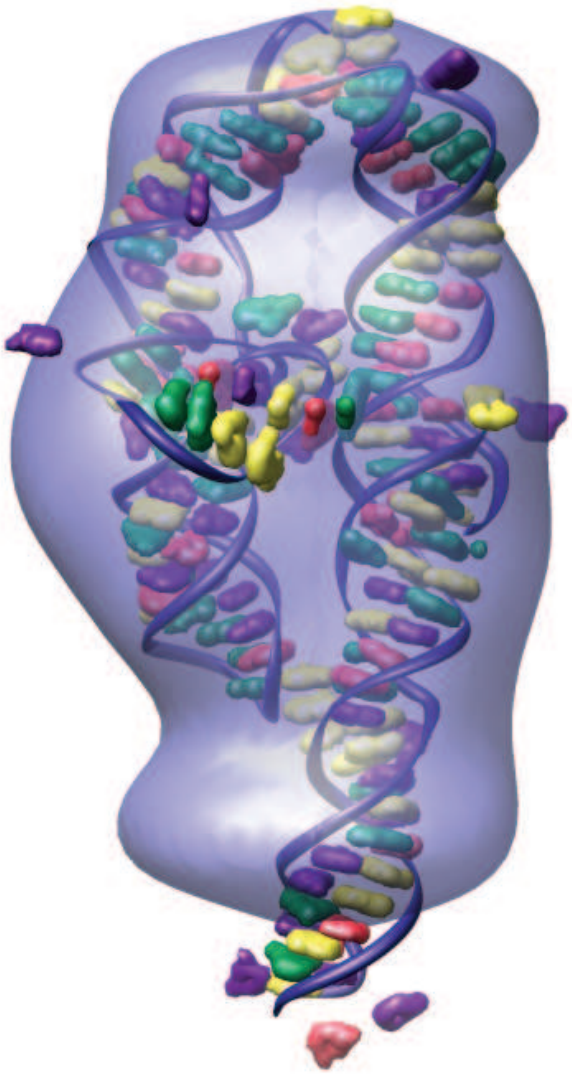

Figure 4.3: Comparison of P4-P6 reconstruction with crystal structure. Reconstruction of folded P4-P6 ($\approx$ 150 ms after folding) is docked with the crystal structure 1GID[21], showing the agreement between the two. Situs[27] was employed to find the best match between the two structures.

rated by a 90° bend. In contrast, this structure is consistent with a roughly equal mixture of initial and final states present in the solution. Thus the reconstruction represents a spatial mean of the ensemble rather than an actual physical state of the molecule. In summary, this series of time resolved reconstructions depicts with large scale accuracy the averaged process of P4-P6 folding.

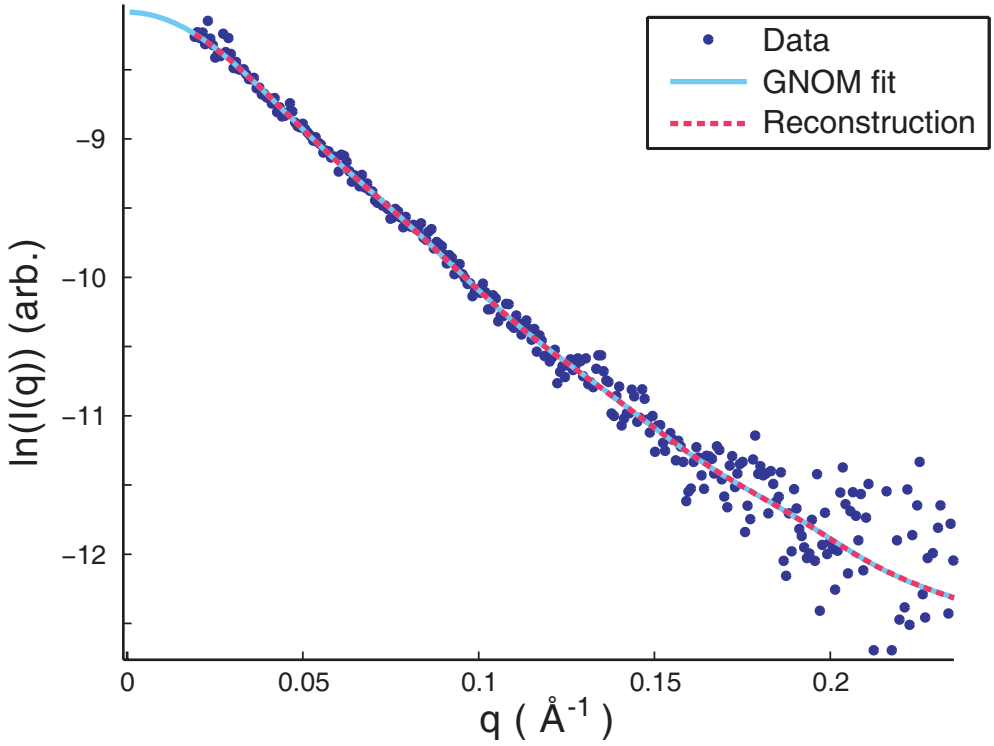

Figure 4.4: GNOM and DAMMIN fits to the experimental data. This plot demonstrates the quality of typical fits to the data by both GNOM and DAMMIN. The dots represent 10.6 ms P4-P6 folding data, which have been scaled by a beam intensity monitor to account for small changes in x-ray intensity. The solid line is the GNOM fit to this data, and the dashed line is calculated from a single reconstruction, both of which are scaled to match the data by the programs which produced them. The curve found by GNOM fits the data well within the noise, but the noise is large compared to the precision of the reconstruction, which leaves room for error. DAMMIN and GNOM curves are indistinguishable, demonstrating how precisely the reconstruction fits the GNOM representation of the data. The consequences of these observations are discussed in the text.

Table 4.1: Calculated S/N for different regions of the 158 ms P4P6 scattering curve. While the S/N is high for most of the curve, in the high q regions where the information about small molecular features is available, the S/N approaches 1. This increases the chances that noise will be misinterpreted as signal in the reconstruction process.

q range (Å^{-1})	S/N
0.019-0.062	39.3
0.063-0.106	40.0
0.107-0.151	15.1
0.151-0.195	5.9
0.196-0.238	1.5

While the P4-P6 domain is a smaller molecule with a relatively simple structure, the full length *Tetrahymena* ribozyme contains many helices to locate. Due to its complex secondary structure, the unfolded state is harder to predict without detailed modeling, and likely corresponds to an ensemble of states. The R_g vs time curve shown in Figure 4.5, is likewise more complex. In this case folding takes several minutes. Three distinct collapse phases are observed under the experimental conditions employed, indicating the existence of two long-lived intermediates in addition to the initial and final states. In an effort to capture and characterize a true time-resolved intermediate state, we reconstructed only points where the R_g of the molecule was not rapidly changing.

The resulting averaged shape envelopes, as well as the associated MNSD values, are displayed in Figure 4.6. According to Volkov and Svergun [29] a MNSD of less than 0.7 indicates ideal agreement between individual reconstructions. In general, we have found this to be true, though when the MNSD is less than 1 strong similarities between different constructs are observed. For this data set, we found the MNSD decreased as folding progressed. Only the final state had an MNSD of less than 0.7.

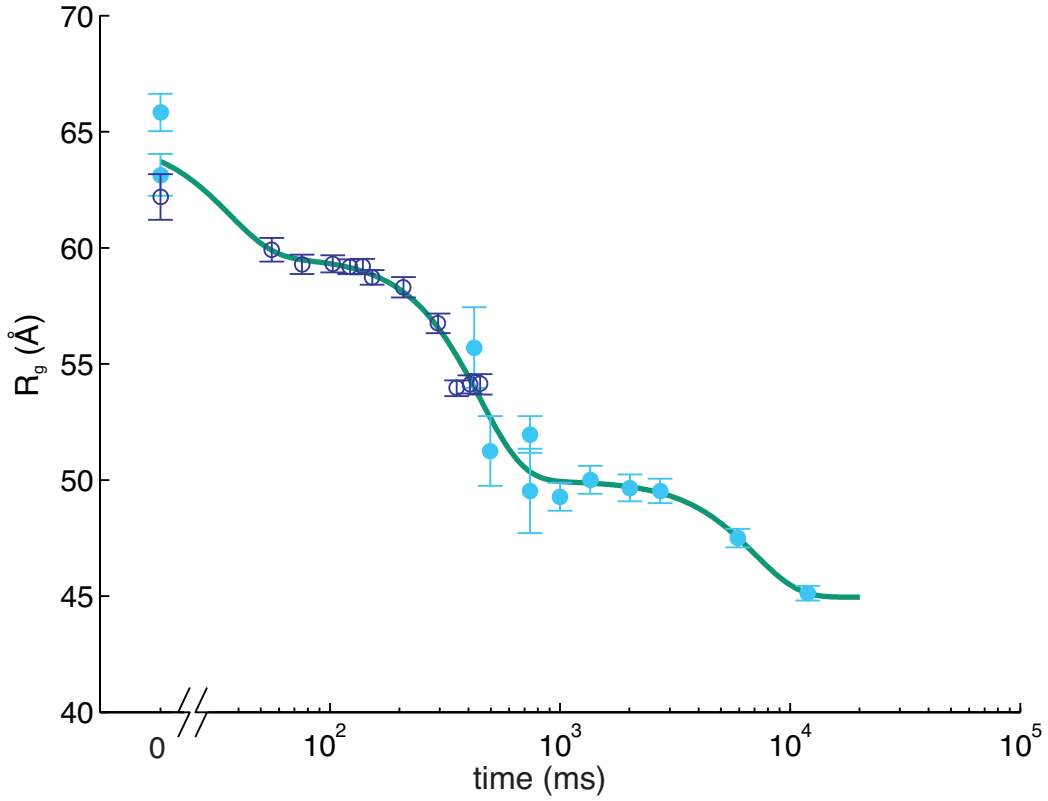

Figure 4.5: Folding time course for the full length ribozyme. The points are calculated from a fit to the Guinier region of the data[2], with the errors being determined by the 95 % confidence intervals. The solid line is a fit to the expression $R_g = (L_0 - L_1)e^{\frac{-t}{\tau_0}} + (L_1 - L_2)e^{\frac{-t}{\tau_1}} + (L_2 - L_3)e^{\frac{-t}{\tau_2}} + L_3$ where $L_0 = 64.7$, $L_1 = 59.6$, $L_2 = 49.98$, $L_3 = 44.95$, $\tau_0 = 0.490$ ms, $\tau_1 = 157.6$ ms and $\tau_2 = 89.1 \times 10^3$ ms. Radius of gyration versus time for the Tetrahymena ribozyme shows two long lived intermediates on the folding pathway. Open circles correspond to data points acquired at the APS with the continuous flow mixer; filled circles represent data collected with a stopped-flow mixer at CHESS.

4.4 Discussion

The first of our initial goals was to determine whether the noise present would prevent accurate reconstruction of time resolved data. Our results address this question. In the case of P4-P6 folding, GNOM produces smooth curves which fit the data reasonably well and reflect the overall differences in the SAXS curves

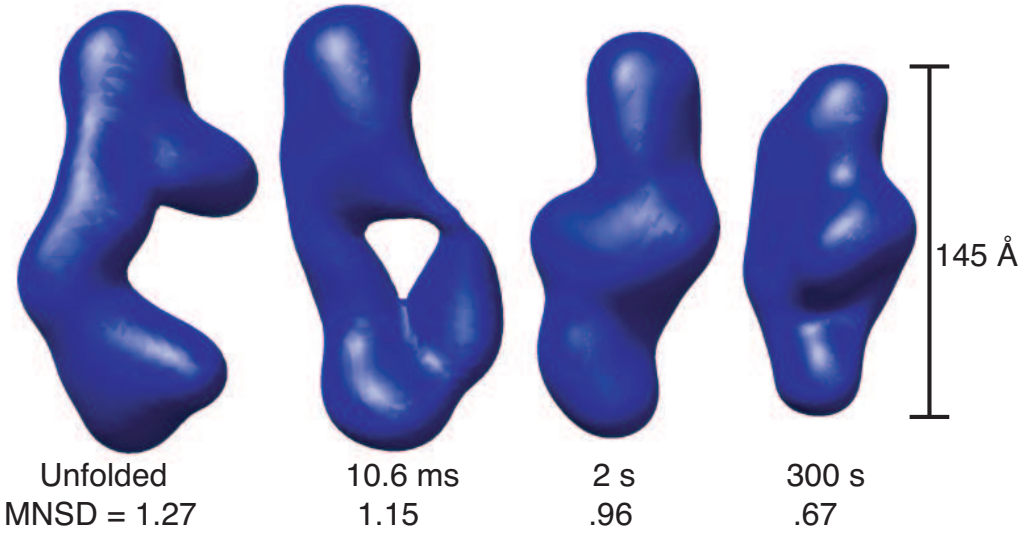

Figure 4.6: Reconstructions for select time points during folding of the full length ribozyme. This figure indicates the time elapsed since mixing and shows the MNSD between ten individual reconstructions of scattering states along the folding pathway of the ribozyme. The initial and final states are shown, along with two states that correspond to the measured plateaus in the curve of R_g vs. time.

measured at different times after folding, despite noise in the data. Reconstructing with DAMMIN then produces shapes consistent on a large scale with data from the crystal structure and models. Variations between the reconstructions on a smaller scale are less reliable. For example, shape envelopes of the 92 and 168 ms time points from P4-P6 folding resemble each other much more closely than that of the intermediate 125 ms data point. Such a switch in an ensemble measurement cannot be physical. Figure 4.7 demonstrates that some of these differences are inherent in the fitted intensity curves. Therefore small deviations may arise from noise in the scattering data since it cannot be fit to the precision with which the GNOM curve is reconstructed.

These inconsistencies might be mitigated by removing some of the highest

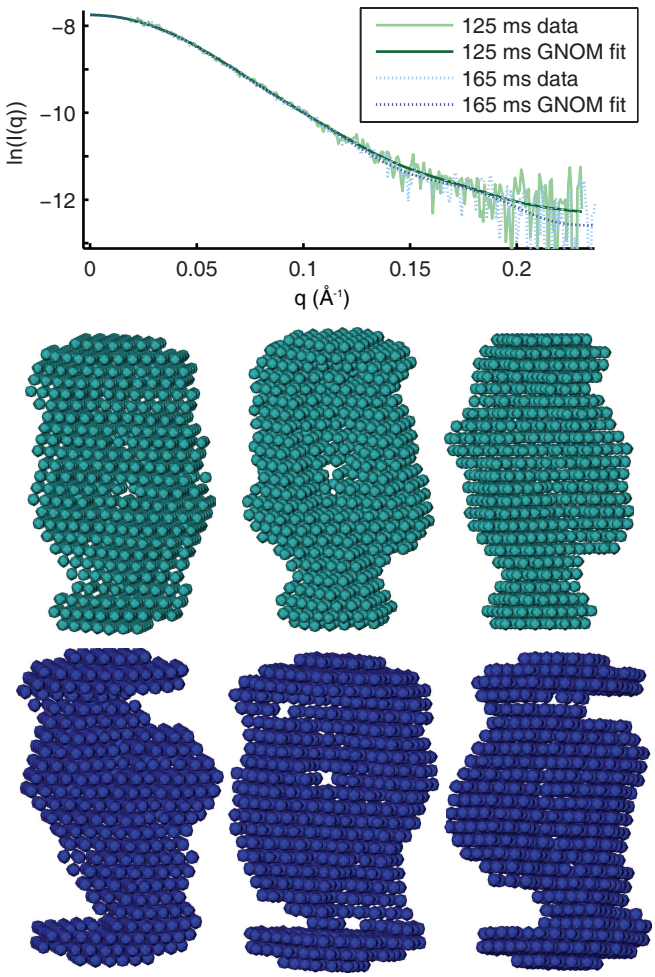

Figure 4.7: Variations in reconstructions of P4-P6. Above is a plot showing differences that arise in the GNOM fit to P4-P6 data acquired at 125 and 168 ms after folding was initiated. The GNOM fit shows variation in the high q region, despite the underlying data being nearly identical within the noise. Below are three individual reconstructions each for the 125 ms (upper) and 168 ms (lower) data. Although there is variation between the models reconstructed from a single curve, the variation is greater between the top and bottom sets of models. This is also reflected quantitatively in the MNSD which is 0.58 and 0.63 for the 125 ms and 168 ms points respectively, but 0.69 between shapes from the different reconstructions, a bit higher. Thus, in spite of overall similarity in the data, models at these two time points reconstruct to slightly different shapes.

q data. In fact, the DAMMIN manual cautions against including the outer portions of all scattering curves, but leaves it to the user to determine appropriate cutoffs, which depend on the details of each experiment. To test the effect of excluding the weak part of the signal, we repeated the reconstruction procedure on the three longest time curves mentioned above after removing the data at $q > 0.18$ Å^{-1} where the output from GNOM showed the greatest variation. The signal to noise ratio for the eliminated data averages to $\approx$2, while the S/N of the remaining scatter increases by approximately 8. Qualitatively, these reconstructions appear to show less detail (supplementary data). However, all three spatial averages had a smaller MNSD than their lower S/N counterparts and we found greater agreement between them. Thus, although some information can be provided by noisy data, there is clearly a point at which the increase in noise leads to inaccuracy. In these cases the data are better removed.

Of all of the states studied, the unfolded ribozyme is the most challenging to reconstruct. Ten reconstructions of this initial scattering curve produced ten different shapes, as evidenced by the large MNSD of the spatial average. Because this complex, unfolded molecule is not mechanically constrained, it is tempting to attribute variations of the reconstructions to supposed heterogeneity of the molecular ensemble. However, the calculated scattering curves for the individual reconstructions agree well with the data in the range of the measurements as shown in Fig 4.8. Therefore, with this experimental resolution, the SAXS data alone do not exclude the possibility that any one of the individual reconstructions might accurately represent the entire ensemble. However, if the scattering profiles of the reconstructed shapes are computed beyond the experimentally accessed range, differences are seen. Initially it may seem counterintuitive that important information exists at high q for a larger molecule, but the extended

nature of the state demands measurements over a broad range of length scales (e.g. to high angle) to describe it completely. Unfortunately, the signal-to-noise is poor at high scattering angles, especially for time-resolved experiments, and as already demonstrated the high q data must contain sufficient S/N to be useful.

While the unfolded state(s) of the ribozyme cannot be reconstructed uniquely from the existing data, we note that the agreement for the ten reconstructions improves for scattering data measured at longer times after mixing, as indicated by the lower MNSD. The drop in R_g indicates that the ribozyme becomes more compact during this time. Although the MNSD is largest for extended states, we note that a simpler extended state (that of P4-P6) reconstructed well. Therefore, and not unexpectedly, reconstructions from states which are complex and extended pose the greatest challenges. Scattering profiles of these molecules should display intensity variations over a very large range of scattering angles. A large MNSD (produced from reconstructions of data acquired in the more accessible q ranges described here) may in fact indicate that the states of interest are both complex and extended.

To address the question of how scattering data from multiple states will affect reconstruction, we turn to the SAXS data from P4-P6 folding which are adequately modeled by a two-state transition. At the time midpoint of the folding transition, where roughly half of the molecules are unfolded and half are folded, the reconstruction converges to an apparent spatial average of the population. This is not entirely unexpected since the SAXS data are a linear combination of scattering curves representing the states present in solution. Insight might be gained by analyzing this curve with SVD[30] and then applying reconstruction

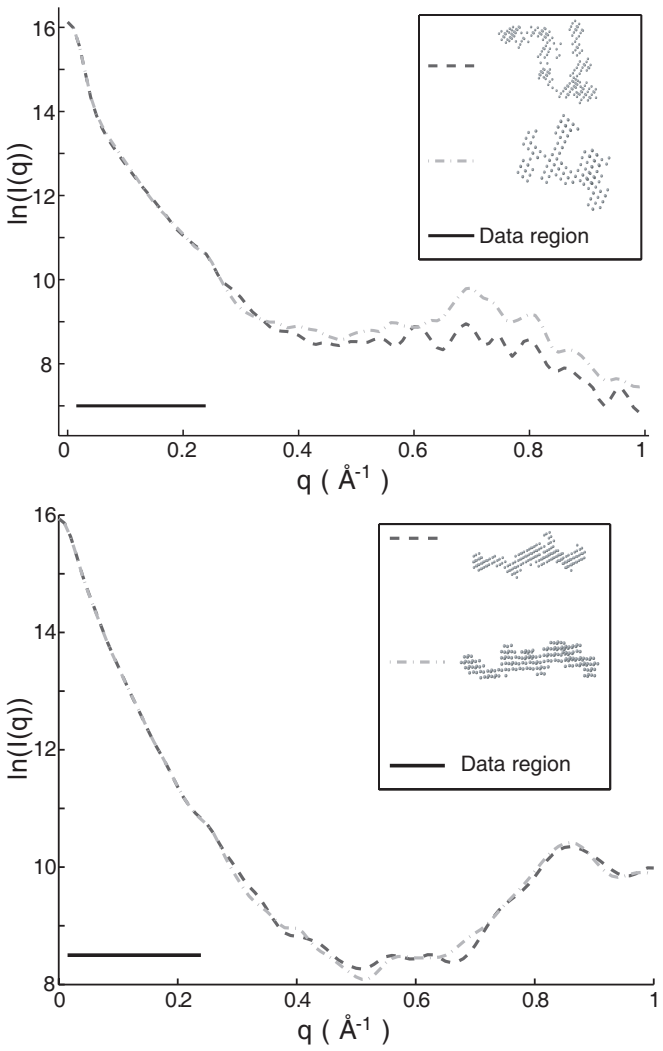

Figure 4.8: Variations between reconstructions from the same SAXS data. This figure shows calculated scattering intensities for the most widely varying reconstructions of the scattering data representing unfolded molecules. The black line in both figures indicates the range over which we have acquired data, from approximately $q = 0.01$ to $q = 0.22$ Å^{-1}. *Top*: Although these reconstructed models for the full length ribozyme are very different, we note that the scattering profiles of both of these models agree well with our measurements (data not shown) and with each other within the data range. Discrepancies which are evident outside of the scattering regime might indicate there is more information to gather at larger q. *Bottom*: For P4-P6 reconstructions the agreement is excellent both within and beyond the data range.

methods, but such an approach generally requires several different scattering curves, a limited number of states in solution, and some additional knowledge of the system to determine the weights of the basis curves needed to build the relevant scattering states. Alternatively, a reconstruction that represents an average rather than a physical state may be employed to analyze bulk properties of the molecule in solution. For example Lipfert et al. [28] demonstrated the utility of reconstructions in electrostatic calculations of nucleic acid solutions. In such a case, the mean shape of a molecule in solution would be an ideal basis for modeling.

4.5 Conclusions

Our results demonstrate the feasibility and utility of reconstructions applied to time resolved data. Even for the most complex molecules, where a unique structure was not readily identifiable, the MNSD may provide an additional measure of the rate of compaction. Furthermore, one could reasonably expect to reconstruct meaningful states near the end of the folding path. For simple molecules with less of a uniqueness problem, intermediate states might be found in this way. Additionally, there are many molecular processes which involve major conformational changes between two compact states with transient intermediates, which may be studied by the methods discussed herein. While scattering data may be noisy, as long as an overall shape of the curve can be elucidated, a reconstruction can accurately represent the prominent dimensions and overall changes of a molecule.

Reconstructions are also useful when comparing new SAXS data to existing

structural data and models. While the scattering curve for a crystal structure can be calculated and compared to the data, differences might be challenging to interpret in the absence of three dimensional shape envelopes. Data from techniques such as electron microscopy, which produces images but not scattering curves are also hard to compare. Finally, there have been recent efforts to combine low resolution data, atomic resolution data, and computer modeling to maximize the information gained from each technique[8;31]. Time resolved reconstructions used in conjunction with these techniques will provide deeper insight into the changes that take place in a molecule.

4.6 Acknowledgement

This work was funded by the Nanobiotechnology Center which is supported in part by the STC Program of the National Science Foundation under Agreement No. ECS-9876771. Additional support was provided by the National Institutes of Health through P01-GM066275 and the National Science Foundation through MCB-0347220. We would like to thank beamline scientists Alec Sandy and Suresh Narayanan for their assistance at APS, and Arthur Woll for his assistance at CHESS. We thank Simon Mochrie for loan of the stopped-flow mixer. Use of the Advanced Photon Source was supported by the U. S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. The Cornell High Energy Synchrotron Source is supported by the National Science Foundation and the National Institutes of Health/National Institute of General Medical Sciences under NSF award DMR-0225180. This work made use of the research computing facility of the Cornell Center for Materials Research (CCMR) with support from the National Science

Foundation Materials Research Science and Engineering Centers (MRSEC) program (DMR 0520404).

REFERENCES

1. Lamb, J. S., L. W. Kwok, X. Qiu, K. Andresen, H. Park, and L. Pollack. "Reconstructing three dimensional shape envelopes from time resolved small angle X-ray scattering data." *J. Appl. Cryst.*, in press.

2. Guinier, A. and G. Fournet. *Small-Angle Scattering of X-Rays*. John Wiley and Sons., Inc., New York, 1955.

3. Svergun, D. "Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing." *Biophys. J.*, **76**(6): 2879–2886, 1999.

4. Walther, D., F. E. Cohen, and S. Doniach. "Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules." *J. Appl. Cryst.*, **33**: 350–363, 2000.

5. Chacón, P., J. F. Díaz, F. Morán, and J. M. Andreu. "Reconstruction of protein form with X-ray solution scattering and a genetic algorithm." *J. Mol. Biol.*, **299**(5): 1289–1302, 2000.

6. Heller, W. T., E. Abusamhadneh, N. Finley, P. R. Rosevear, and J. Trewhella. "The solution structure of a cardiac troponin C-troponin I-troponin T complex shows a somewhat compact troponin c interacting with an extended troponin I-Troponin T component." *Biochemistry*, **41**(52): 15654–15663, 2002.

7. Svergun, D. I. and M. H. J. Koch. "Advances in structure analysis using small-angle scattering in solution." *Curr. Opin. Struct. Biol.*, **12**(5): 654–660, 2002.

8. Grossmann, J. G. "Biological solution scattering: recent achievements and future challenges." *J. Appl. Cryst.*, **40**: S217–S222, 2007.

9. Nöllmann, M., W. M. Stark, and O. Byron. "Low-resolution reconstruction of a synthetic DNA Holliday junction." *Biophys. J.*, **86**(5): 3060–3069, 2004.

10. Funari, S. S., G. Rapp, M. Perbandt, K. Dierks, M. Vallazza, C. Betzel, V. A. Erdmann, and D. I. Svergun. "Structure of free *Thermus flavus* 5 S rRNA at 1.3 nm resolution from synchrotron x-ray solution scattering." *J. Biol. Chem.*, **275**(40): 31283–31288, 2000.

11. Moody, M. F., P. Vachette, A. M. Foote, A. Tardieu, M. H. Koch, and J. Bordas. "Stopped-flow x-ray scattering: the dissociation of aspartate transcarbamylase." *Proc. Natl. Acad. Sci. U. S. A.*, **77**(7): 4040–3, 1980.

12. Tsuruta, H., T. Nagamura, K. Kimura, Y. Igarashi, A. Kajita, Z. X. Wang, K. Wakabayashi, Y. Amemiya, and H. Kihara. "Stopped-Flow apparatus for X-ray-scattering at subzero temperature." *Rev. Sci. Instrum.*, **60**(7): 2356–2358, 1989.

13. Pollack, L., M. W. Tate, N. C. Darnton, J. B. Knight, S. M. Gruner, W. A. Eaton, and R. A. Austin. "Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **96**: 10115–10117, 1999.

14. Fang, X. W., K. Littrell, X. Yang, S. J. Henderson, S. Siefert, P. Thiyagarajan, T. Pan, and T. R. Sosnick. "Mg^{2+}-dependent compaction and folding of yeast tRNA(Phe) and the catalytic domain of the *B-subtilis* RNase P RNA determined by small-angle X-ray scattering." *Biochemistry*, **39**(36): 11107–11113, 2000.

15. Russell, R., I. S. Millett, M. W. Tate, L. W. Kwok, B. Nakatani, S. M. Gruner, S. G. J. Mochrie, V. P., S. D., D. Herschlag, and L. Pollack. "Rapid compaction during RNA folding." *Proc. Natl. Acad. Sci. U.S.A.*, **99**(7): 4266–4271, 2002.

16. Akiyama, S., S. Takahashi, T. Kimura, K. Ishimori, I. Morishima, Y. Nishikawa, and T. Fujisawa. "Conformational landscape of cytochrome *c* folding studied by microsecond-resolved small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **99**: 1329–1334, 2002.

17. Schlatterer, J. C., L. W. Kwok, J. S. Lamb, H. Y. Park, K. Andresen, M. Brenowitz, and L. Pollack. "Hinge stiffness is a barrier to RNA folding." *J. Mol. Biol.*, **379**(4): 859–870, 2008.

18. Das, R., L. W. Kwok, I. S. Millett, Y. Bai, T. T. Mills, J. Jacob, G. S. Maskel, S. Seifert, S. G. J. Mochrie, P. Thiyagarajan, S. Doniach, L. Pollack, and D. Herschlag. "The fastest global events in RNA folding: Electrostatic relaxation and tertiary collapse of the tetrahymena ribozyme." *J. Mol. Biol.*, **332**(2): 311–319, 2003.

19. Pan, J., D. Thirumalai, and S. A. Woodson. "Magnesium-dependent folding of self-splicing RNA: Exploring the link between cooperativity, thermodynamics, and kinetics." *Proc. Natl. Acad. Sci. U. S. A.*, **96**(11): 6149–6154, 1999.

20. Rook, M. S., D. K. Treiber, and J. R. Williamson. "An optimal Mg^{2+} concentration for kinetic folding of the *Tetrahymena* ribozyme." *Proc. Natl. Acad. Sci. U. S. A.*, **96**(22): 12471–12476, 1999.

21. Cate, J. H., A. R. Gooding, E. Podell, K. H. Zhou, B. L. Golden, C. E. Kundrot, T. R. Cech, and J. A. Doudna. "Crystal structure of a group I ribozyme domain: Principles of RNA packing." *Science*, **273**(5282): 1678–1685, 1996.

22. Guo, F., A. R. Gooding, and T. R. Cech. "Structure of the Tetrahymena ribozyme: Base triple sandwich and metal ion at the active site." *Molecular Cell*, **16**(3): 351–362, 2004.

23. Kwok, L. W., I. Shcherbakova, J. S. Lamb, H. Y. Park, K. Andresen, H. Smith, M. Brenowitz, and L. Pollack. "Concordant exploration of the kinetics of RNA folding from global and local perspectives." *J. Mol. Biol.*, **355**(2): 282–293, 2006.

24. Russell, R. and D. Herschlag. "New pathways in folding of the Tetrahymena group I RNA enzyme." *J. Mol. Biol.*, **291**(5): 1155–1167, 1999.

25. Konarev, P. V., M. V. Petoukhov, V. V. Volkov, and D. I. Svergun. "ATSAS 2.1, a program package for small-angle scattering data analysis." *J. Appl. Cryst.*, **39**: 277–286, 2006.

26. Svergun, D. I. "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." *J. Appl. Cryst.*, **25**: 495–503, 1992.

27. Wriggers, W. and P. Chacón. "Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering." *J. Appl. Cryst.*, **34**: 773–776, 2001.

28. Lipfert, J., L. Columbus, V. B. Chu, and S. Doniach. "Analysis of small-angle X-ray scattering data of protein-detergent complexes by singular value decomposition." *J. Appl. Cryst.*, **40**: S235–S239, 2007.

29. Volkov, V. V. and D. I. Svergun. "Uniqueness of ab initio shape determination in small-angle scattering." *J. Appl. Cryst.*, **36**: 860–864, 2003.

30. Chen, L., K. O. Hodgson, and S. Doniach. "A lysozyme folding intermediate revealed by solution X-ray scattering." *J. Mol. Biol.*, **261**(5): 658–671, 1996.

31. Suhre, K., J. Navaza, and Y. H. Sanejouand. "*NORMA*: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps." *Acta Cryst., Sect. D: Biol. Cryst.*, **62**: 1098–1100, 2006.

CHAPTER 5

TIME-RESOLVED DIMERIZATION OF A PAS-LOV PROTEIN MEASURED WITH PHOTOCOUPLED SMALL ANGLE X-RAY SCATTERING

The flow cell described in Chapter 2 was employed to collect measurements on the conformation changes which accompanied blue-light absorption by a PAS (Per Arnt Sim) domain protein. This chapter describes initial work on this project, and is reproduced with minimal alteration from a communication in the *Journal of the American Chemical Society*[1].

Here, we report the coupling of photoexcitation with time-resolved SAXS to monitor conformational changes accompanying light activation of the blue-light signal transduction protein Vivid (VVD). Solution small angle X-ray scattering (SAXS) reports the size and shape of soluble biomolecules. Previously, SAXS was used in conjunction with rapid mixing techniques to time resolve macromolecular folding[2;3]. SAXS studies have also been used to determine large scale differences between the dark and light excited states of proteins[4;5]. Flash-flow devices have been used for time-resolved IR spectroscopy[6]; integration with SAXS expands the technique, enabling exploration of global kinetics. While time dependent processes in different blue-light sensors have been revealed by a variety of techniques[7;8], this method elucidates previously unknown association dynamics of the long-lived VVD light-adapted state.

VVD, a so-called LOV (for *L*ight *O*xygen *V*oltage sensing) protein of the PAS family, regulates blue-light responses in the filamentous fungus *Neurospora crassa*[9]. Photon absorption by the VVD flavin cofactor drives conformational changes within the LOV domain. In the absence of light, the protein

is monomeric. Recent measurements suggest the establishment of a rapidly exchanging monomer:dimer equilibrium in the light-activated state[10], thus VVD (like other LOV domains[11;12]) changes association state in response to light-stimulated structural modifications. PAS:PAS dimerization is believed to be a key regulatory event in signal transduction[13], and is likely important for a close VVD homologue, WC-1, to activate transcription[14]. Because of the challenge of characterizing structural intermediates in these processes, little is known about how cofactor chemical state relates to association mode, and ultimately, the engagement of targets.

To time-resolve changes in the structure of VVD resulting from photoexcitation, we employed a microfluidic cell coupled to both laser and X-ray sources (Figure 5.1). The cell consists of a thin wall polyester tube (Advanced Polymers, VT)[15] which is optically transparent and scatters X-rays minimally. A 473 nm laser beam (Holograms & Lasers International, Texas) is directed at 90° to a focused X-ray beam. The location where the X-ray and light beams intersect defines the time-equals-zero position for the experiment. Offsetting the two beams creates a controllable delay between laser excitation and SAXS measurement. This average delay time is calculated by dividing the distance between the two beams, x, by the flow speed of the protein solution, u (Figure 5.1). Use of a flow cell eliminates radiation damage to the protein and enables long X-ray exposures, which increase signal-to-noise without loss of time resolution. Temporal precision is restricted by the laminar flow boundary condition at the channel wall where u=0, but can be improved by employing a flow of sheath buffer which confines the protein-containing solution to a thin jet that travels along the tube axis with relatively uniform speed. The transit time of each molecule through either the laser or X-ray incident spot was 160 ms on average. The laser

spot was continuously illuminated for all light-state measurements and the X-ray exposure time ranged from 20 to 30 sec.

The intensity of the scattered X-rays is typically plotted as a function of the momentum transfer $q = 4\pi\frac{\sin(\theta)}{\lambda}$, where θ is half the scattering angle and λ is the X-ray wavelength. Following standard analysis procedures[16] the zero angle intensity, $I(0)$, and the radius of gyration, (R_gg), were extracted from the low q data (see Supporting Information). A Kratky plot[17] places emphasis on data acquired at large q, which provides detail about macromolecular conformation.

SAXS measurements of the dark state of VVD are in good agreement with the computed signal from the monomeric crystal structure 2PD7 (Supporting Information)[5]. Upon light-illumination, the low angle scattering intensity from VVD is increased relative to the dark state (Figure 5.2 and Supporting Information). The increase occurs rapidly, and appears finished by the earliest time detected, 20 ms after photoexcitation. Such an increase in $I(0)$ is consistent with dimerization. Inline multi-light scattering (MALS), dynamic light scattering (DLS), size exclusion chromatography and equilibrium ultracentrifugation confirm the presence of a rapidly dissociating dimer in the light adapted state[10]. Complete dimerization in 20 msec is consistent with a diffusion controlled association rate constant in the range of $10^5 - 10^6$ M^{-1} s^{-1} [10], and demonstrates that the conformational change which precedes dimerization is unlikely to be rate limiting. This association rate exceeds that reported for other LOV domain proteins[12]. Importantly, we also evaluated the time-dependent SAXS profile of a VVD point mutant (Cys71Ser) that cannot undergo light-induced dimerization, but otherwise has normal photochemical properties4. Cys71Ser VVD undergoes no change in scattering upon illumination; thus, laser heating or ra-

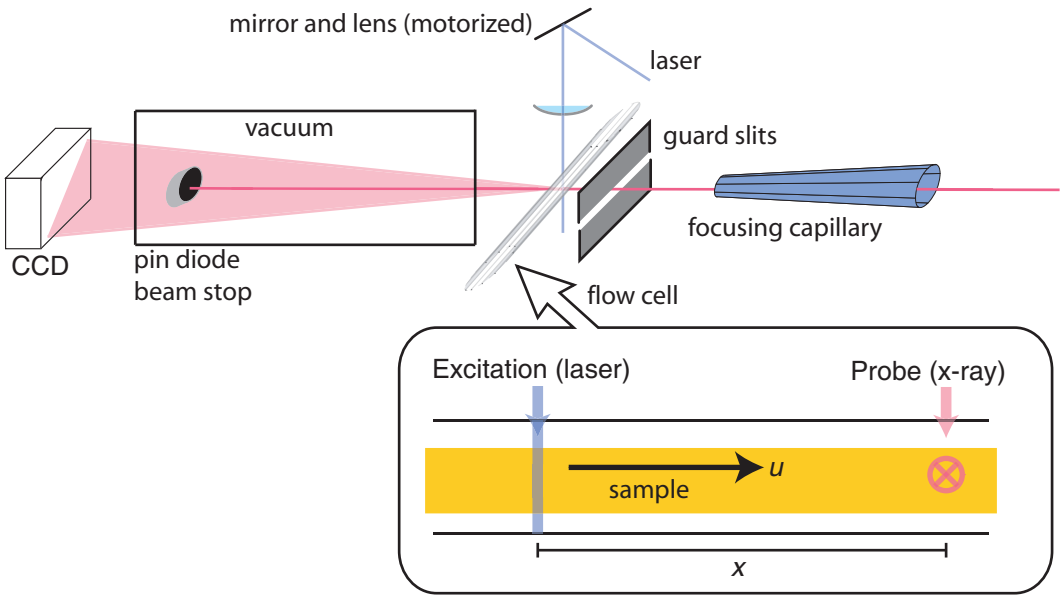

Figure 5.1: A schematic of the continuous flow setup that enables time-resolved SAXS measurements following photoexcitation of protein. An X-ray beam, incident from the right, passes through a focusing capillary15,16 to obtain sufficient X-ray flux in a small spot. Guard slits are used to decrease the scattering background. The excitation laser beam is directed perpendicular to the axis of the flow cell and the X-ray beam. A magnified illustration of the flow cell demonstrates how time-resolution is achieved.

diation damage cannot be the cause of the scattering changes we observe with wild-type VVD.

Scattering data acquired at longer times after photoexcitation show variation in the high q region of the profile, suggesting that the conformational changes continue for several seconds after dimerization is complete (Figure 5.3). As VVD remains in the light-adapted state for several hours[18] this phenomenon reflects additional changes post dimerization that may be essential for engaging targets and propagating signals.

While these time-resolved SAXS measurements show a consistent pattern of oligomerization after light excitation, the magnitude of the change varies depending on the specific sample and experiment. The affinity of the VVD dimer

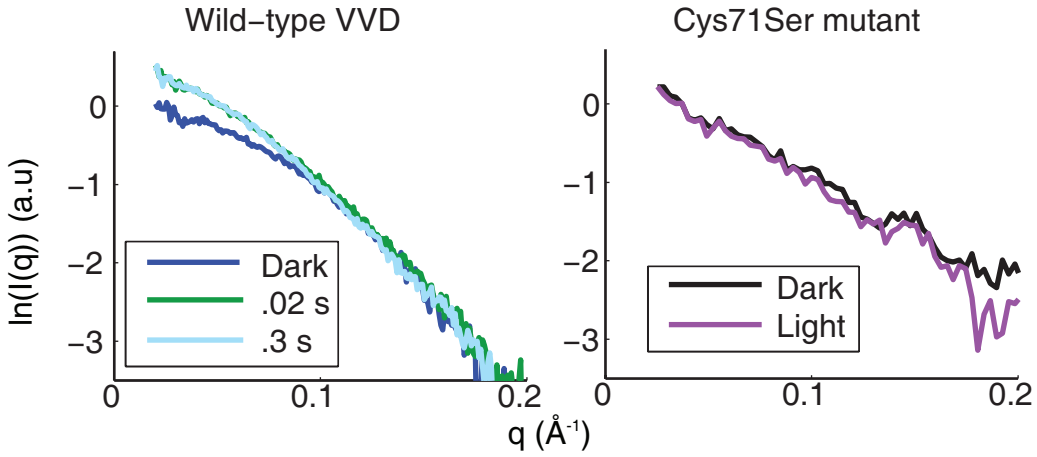

Figure 5.2: Left: Scattering profiles for wild-type VVD before (blue) and 0.02 s after excitation (green) display an increase in zero angle intensity, consistent with light induced dimerization. No further change is measured 0.3 s after excitation (cyan). Right: Scattering profiles of the inactive Cys71Ser mutant are not changed by light excitation.

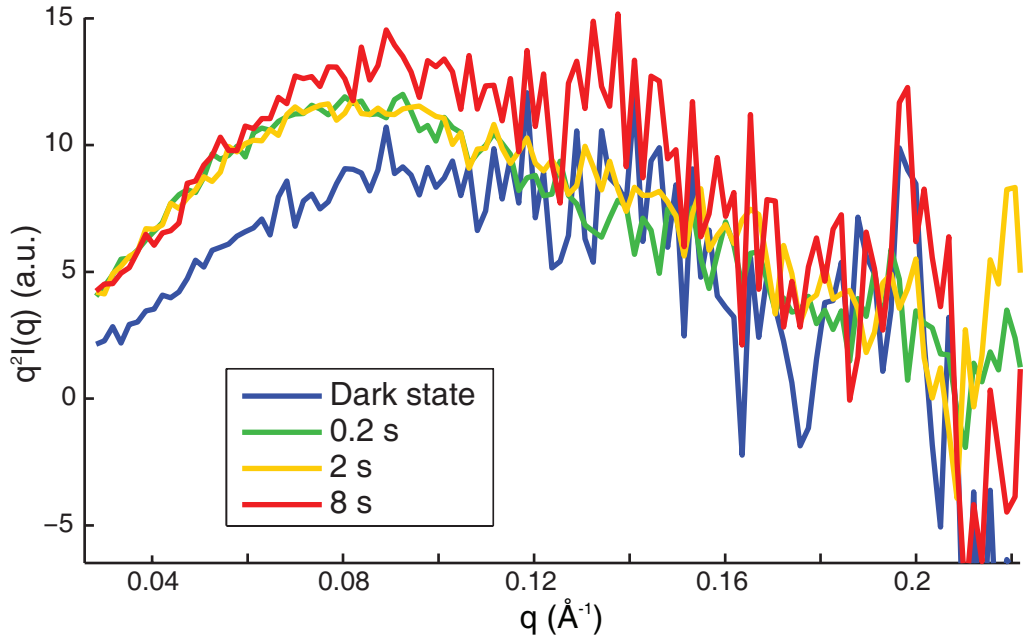

Figure 5.3: Kratky plots of data collected hundreds of milliseconds to seconds after photoexcitation. Variations in scattering profiles become evident at q¿0.08 Å^{-1} after 8 s indicating conformational changes of the dimer.

is known to be highly sensitive to modifications at the N-terminus as well as protein oxidation at currently unspecified sites[10]. We suspect differences in dimer yield result from such subtle changes in protein chemistry; we are exploring the phenomenon in more detail. Measurements on millisecond (or submillisecond) time scales will be required to directly detect light-induced conformational differences in the monomeric state which must precede and gate access to the dimeric state

Time-resolved measurements of global structural states are critical for unraveling the protein dynamics that underlie LOV protein light sensing; SAXS is an ideal tool for studying both the conformational changes and protein association that accompany such events. Application of a continuous flow cell to couple SAXS and photoexcitation, as demonstrated here, can elucidate millisecond conformational changes without signal averaging by photocycle repetition, thus avoiding the sample damage associated with repeat exposures and facilitating measurement of samples with slow photocycles. Notably, this method will resolve diffusion-limited association rates and it is generally applicable to many photoreceptors, even those with transient light-adapted states much less stable than that of VVD.

5.1 Acknowledgement

We thank Arthur Woll, Kurt Andresen, Xiangyun Qiu, Li Li, and Sterling Cornaby for help with data collection. This work was supported by the Cornell Nanobiotechnology Center (NSF ECS-9876771) NSF grant MCB-0347220 NIH

grant GM079679. This research was conducted at the Cornell High Energy Synchrotron Source (CHESS) which is supported by the NSF and the NIH..

REFERENCES

1. Lamb, J., B. D. Zoltowski, S. A. Pabit, B. R. Crane, and L. Pollack. "Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering." *J. Am. Chem. Soc.*, 2008. Doi:10.1021/ja804236f.

2. Moody, M. F., P. Vachette, A. M. Foote, A. Tardieu, M. H. Koch, and J. Bordas. "Stopped-flow x-ray scattering: the dissociation of aspartate transcarbamylase." *Proc. Natl. Acad. Sci. U. S. A.*, **77**(7): 4040–3, 1980.

3. Pollack, L., M. W. Tate, N. C. Darnton, J. B. Knight, S. M. Gruner, W. A. Eaton, and R. A. Austin. "Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **96**: 10115–10117, 1999.

4. Nakasako, M., T. Iwata, D. Matsuoka, and S. Tokutomi. "Light-induced structural changes of LOV domain-containing polypeptides from Arabidopsis phototropin 1 and 2 studied by small-angle X-ray scattering." *Biochemistry*, **43**(47): 14881–90, 2004.

5. Zoltowski, B. D., C. Schwerdtfeger, J. Widom, J. J. Loros, A. M. Bilwes, J. C. Dunlap, and B. R. Crane. "Conformational switching in the fungal light sensor Vivid." *Science*, **316**(5827): 1054–7, 2007.

6. Toepke, M. W., S. H. Brewer, D. M. Vu, K. D. Rector, J. E. Morgan, R. B. Gennis, P. J. A. Kenis, and R. B. Dyer. "Microfluidic flow-flash: Method for investigating protein dynamics." *Anal. Chem.*, **79**(1): 122–128, 2007.

7. Harper, S. M., L. C. Neil, I. J. Day, P. J. Hore, and K. H. Gardner. "Conformational changes in a photosensory LOV domain monitored by time-resolved NMR spectroscopy." *J. Am. Chem. Soc.*, **126**(11): 3390–1, 2004.

8. Kennis, J., I. van Stokkum, S. Crosson, and M. Gauden. "The LOV2 domain of phototropin: a reversible photochromic switch." *J. Am. Chem. Soc*, 2004.

9. Loros, J. J. and J. C. Dunlap. "Genetic and molecular analysis of circadian rhythms in Neurospora." *Annu. Rev. Physiol.*, **63**: 757–94, 2001.

10. Zoltowski, B. and B. Crane. "Light activation of the LOV protein vivid generates a rapidly exchanging dimer." *Biochemistry*, **47**(27): 7012–7019, 2008.

11. Harper, S. M., L. C. Neil, and K. H. Gardner. "Structural basis of a phototropin light switch." *Science*, **301**(5639): 1541–4, 2003.

12. Nakasone, Y. "Kinetic Measurement of Transient Dimerization and Dissociation Reactions of Arabidopsis Phototropin 1 LOV2 Domain." *Biophysical Journal*, **91**(2): 645–653, 2006.

13. Card, P. B., P. J. A. Erbel, and K. H. Gardner. "Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization." *Journal of Molecular Biology*, **353**(3): 664–77, 2005.

14. Froehlich, A. C., Y. Liu, J. J. Loros, and J. C. Dunlap. "White Collar-1, a circadian blue light photoreceptor, binding to the frequency promoter." *Science*, **297**(5582): 815–9, 2002.

15. Kalinin, Y., J. Kmetko, A. Bartnik, A. Stewart, R. Gillilan, E. Lobkovsky, and R. Thorne. "A new sample mounting technique for room-temperature macromolecular crystallography." *J. Appl. Cryst.*, **38**(2): 333–339, 2005.

16. Guinier, A. and G. Fournet. *Small-Angle Scattering of X-Rays*. John Wiley and Sons., Inc., New York, 1955.

17. Glatter, O. and O. Kratky, editors. *Small Angle X-ray Scattering*. Academic Press Inc., London, 1982.

18. Schwerdtfeger, C. and H. Linden. "VIVID is a flavoprotein and serves as a fungal blue light photoreceptor for photoadaptation." *EMBO J*, **22**(18): 4846–55, 2003.

CHAPTER 6

ILLUMINATING CONFORMATIONAL CHANGES OF A LOV-DOMAIN PROTEIN WITH PHOTOCOUPLED SMALL ANGLE X-RAY SCATTERING

Investigations into the conformational changes of VVD continued to prove fruitful. Additional SAXS studies observed the presence of a second population of dark-state molecules. Careful analysis of this data, as well as study of a variant meant to mimic the light-excited state seen in the crystal structure, reveal new insights into the conformational changes that govern the light-excited dimerization of VVD. This chapter is taken from a manuscript which is in preparation for publication[1].

6.1 Introduction

An organism's survival depends on its ability to sense and react to its environment. The PAS (Per Arnt Sim) superfamily is comprised of modular proteins that respond to a diverse set of local stimuli. The varied functions of PAS family proteins include kinase activation, clock function and ion channel regulation[2]. A structurally conserved subdomain capable of transducing a particular input signal characterizes these proteins. PAS-LOV (light oxygen and voltage) proteins use a sensory ligand to monitor factors important for redox chemistry. For example, the absorption of a blue photon by a flavin ligand can trigger conformational or chemical changes that regulate cellular behavior[3]. Plant phototropins autophosphorylate to control growth[4;5], white collar 1 (WC1) regulates the circadian rhythm in fungi[6], and YtvA modulates the stress response in bacteria[7], all in response to light.

Here, we consider the light-activated response of the PAS-LOV photosensor Vivid, by monitoring global changes in protein structure triggered by the absorption of a single photon. Vivid (VVD), which is found in the fungus *neurospora crassa*, regulates carotenoid production in response to blue light and is implicated in circadian clock function[8]. Crystallographic studies[9] revealed the formation of a cysteine adduct in response to light excitation. In both the dark- and light-state crystal form, VVD is a dimer. In solution, biochemical assays indicate that the protein is monomeric, and upon light activation, undergoes larger conformational changes than seen in the crystal and dimerizes[10;11]. Here, we identify the conformational changes that comprise this important regulatory step.

PAS:PAS dimerization is not unique to VVD and plays a significant role in signaling in other PAS domains[5;6]. Whereas stable PAS dimers have been structurally characterized[12–14], their solution dynamics and *in vivo* function remain elusive. Solution small angle X-ray scattering (SAXS) reports global structural characteristics of proteins and is ideal for studying large conformational changes, including association[15]. Using SAXS we resolve important structural characteristics of the light- and dark-adapted states of two VVD constructs: wildtype VVD-36 and the C71V:C183S variant designed to structurally mimic the light-activated state. We employ a flow cell which enables time resolved measurements as rapidly as 20 ms after photoexcitation, and furthermore, mitigates radiation damage from X-rays which can confound the interpretation of SAXS data. Two distinct dark states as well as a unique dimer were discovered and analyzed. In conjunction with reconstructions of the SAXS data into low resolution shapes, we demonstrate how the regulation of conformational dynamics gates dimerization. This regulatory gate is likely employed to achieve

transient dimerization in other LOV/PAS systems.

6.2 Results

6.2.1 VVD Wildtype

Time dependent SAXS data were collected from multiple batches of protein prepared for these studies (see Materials and Methods). Several batches were required due to the substantial sample consumption requirements. Comparison of data collected from different batches revealed significant variations between VVD monomers. VVD Dark state scattering

SAXS profiles collected in the absence of light excitation consistently reveal two distinct conformations of the dark-adapted protein, shown in Figure 6.1. The structure of the monomer was conserved within each batch of protein: no interconversion was measured on a time scale of 24 hours. The radius of gyration, R_g, calculated from each profile was 18.1 ± 1.1 Å ($\text{Dark}_{\text{crys}}$) and 18.8 ± 0.6 Å (Dark_{alt}). The latter value is larger, although the difference is within the error of the measurement.

To further elucidate the differences between these conformations, the data were compared to theoretical scattering curves calculated from a monomeric subunit of the (dimeric) dark-state crystal structures (2PD7.pdb)[9]. Excellent agreement was found with one of the two dark state curves (Figure 6.1), which we will refer to as $\text{Dark}_{\text{crys}}$. The other, Dark_{alt}, exhibits higher scattering intensity in the range $q > .18$ Å^{-1}, suggesting a more extended structure. As

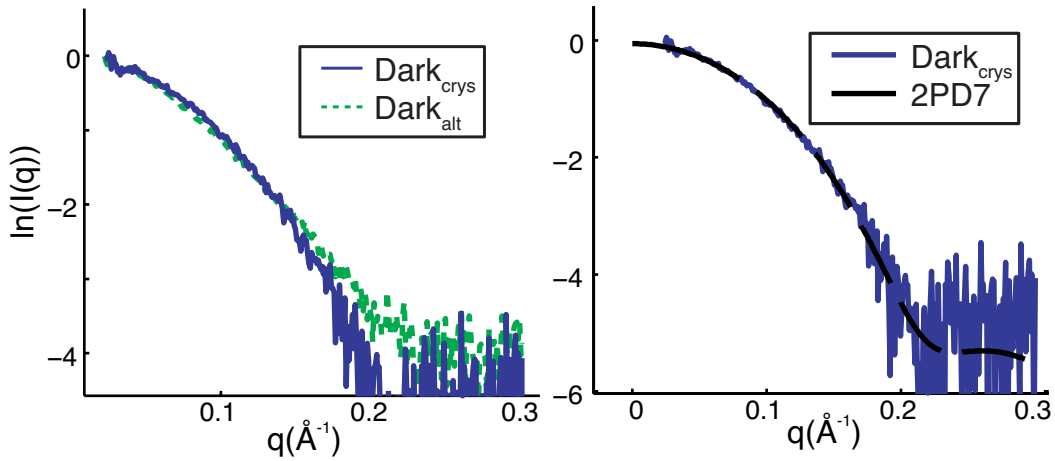

Figure 6.1: The left plot shows two different scattering profiles of the dark state conformation of VVD. The dotted profile shows "Dark_{alt}" state named in the text. The "$\text{Dark}_{\text{crys}}$" scattering is depicted in blue on both the left and the right; on the right it is compared to the computed scattering from PDB structure of the VVD monomer, shown in black[9]. The fit of this scattering curve to the model gives $\chi^2 = .548$.

discussed below, this scattering profile agrees with a structural model proposed for a carefully selected variant. Further analysis confirms this interpretation and provides insight into conformational gating of VVD function.

Calculation of the distance distribution functions, P(r) assists interpretation of subtle differences in scattering profiles, shown in Figure 6.2. The maximum molecular dimension, D_{max}, needed to best reproduce the scattering data was larger for Dark_{alt} than for $\text{Dark}_{\text{crys}}$. Moreover, the slightly higher values of P(r) computed for Dark_{alt} at short length scales are consistent with the undocking of a narrow peptide chain such as a single loop or α-helix. Interestingly, recent crystal structures of VVD-36 demonstrate the presence of an alternate conformation of the first eight N-terminal residues. Furthermore, conformational changes prior to light-state dimerization are believed to involve the N-terminus[10]. With this in mind, modeling was used to assess whether reorganization of residues near the N-terminus of the protein could account for the measured differences.

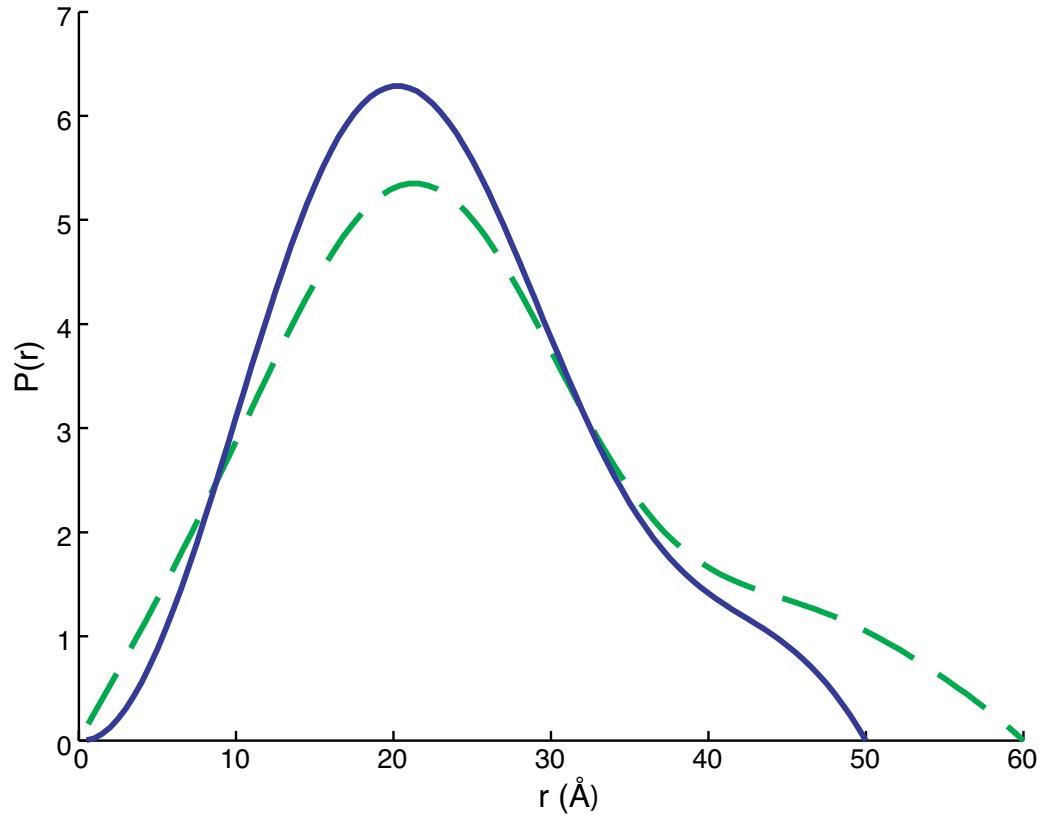

Figure 6.2: P(r) profiles for both VVD wildtype dark states. The solid profile is the inversion of the scattering from the $Dark_{crys}$ scattering profile; the dashed profile corresponds to the $Dark_{alt}$ state. The $Dark_{alt}$ conformation has a larger apparent maximum dimension, and a greater proportion of long dimensions, indicating a more extended molecule. Additional population density at short dimensions indicates additional narrow portions in the molecular structure.

Computational studies were performed to match the SAXS data to structures with varying numbers of N-terminal domain residues undocked. The program EOM (`http://www.embl-hamburg.de/ExternalInfo/Research/Sax/software.html`) generates an ensemble of unfolded or partly unfolded structures from an amino acid sequence and, with a genetic algorithm, searches for a subset of these that combine to match the experimental data. This analysis was performed six times, including selectively longer portions of the N terminal helix in the unfolded ensemble (Figure 6.3). Notably, extension of only the terminal eight peptides, as proposed from the alternate crystal structure, is in-

sufficient to describe the data. At a minimum, the N-terminal peptide chain and the N-terminal α-helix must be displaced to obtain good agreement between the calculated and measured SAXS profiles. Thus, while the $\text{Dark}_{\text{crys}}$ data matches the crystal structure, the Dark_{alt} curve can only be simulated by extending at least 22 amino acids from this structure.

6.2.2 VVD light state data

To elucidate the form of the light-activated dimer, SAXS data were collected 8.8 s after photoactivation of VVD. Experiments were initiated from both of the dark states described above. The most pronounced difference resulting from light-activation, is the dramatic increase in scattering intensity at low q, seen in Figure 6.4, which is consistent with previous observations[11]. Following standard Guinier analysis, the zero-angle intensity is measured and found to increase by factors of 1.76 ± 0.07 and 1.64 ± 0.04 with $\text{Dark}_{\text{crys}}$ and Dark_{alt} as respective initial states. These changes are consistent with rapidly interconverting light-state dimers identified in VVD[10]. However, this increase is smaller than the factor of 2 expected if 100 % of the sample converted to dimers. Spectroscopic measurements (data not shown) reveal only 70 % of the population forms the photo-adduct under these solution and excitation conditions. Thus, the scattering collected from the light-excited states is an admixure of at least two states for each curve: light-activated dimers and dark state monomers.

Interestingly, the light-excited SAXS profiles preserve the shape difference observed in the high q range of the dark-state monomer profiles. These discrepancies in the scattering data could reflect different light-state dimers originating

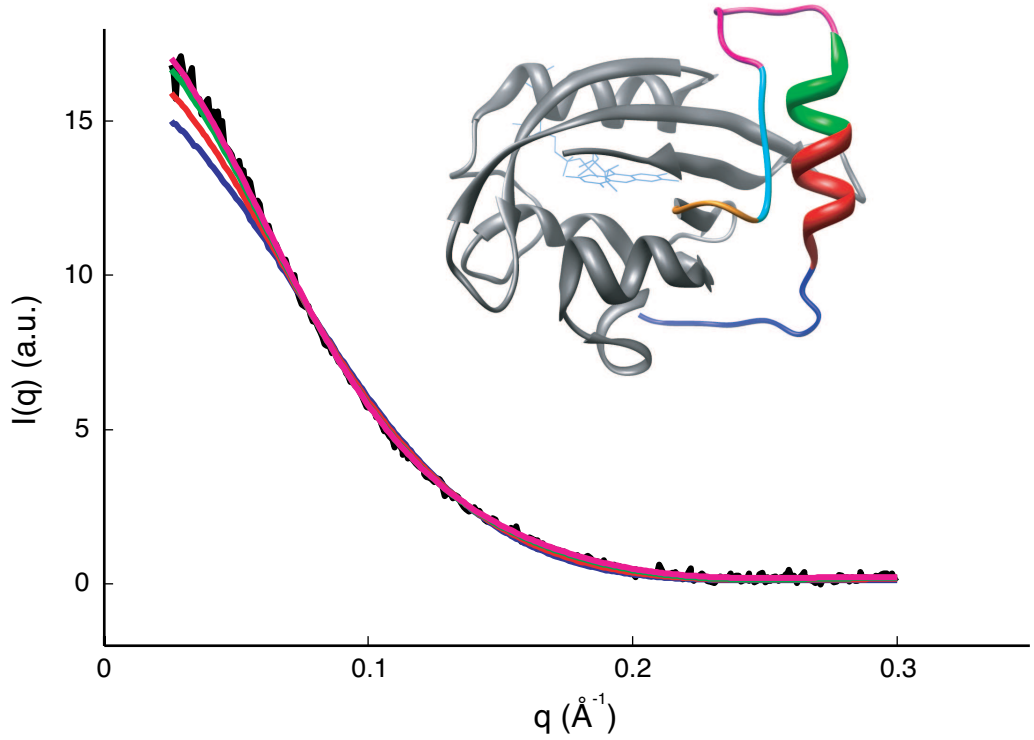

Figure 6.3: Results of the EOM analysis. The smooth curves indicate attempts to fit to the data with an ensemble of partially extended molecules. They represent models in which varying lengths of the amino acid sequence are flexible. From lowest to highest, 11, 18, 22, and 29 residues are displaced. The scattering curves generated from the first two ensembles do not fit the data well. Extending residues 36-58, which includes the entire N-terminal α-helix, produces a reasonable fit. Adding the loop that connects the α-helix with the LOV domain improves the match to the data. Inset is the 3D72 crystal structure. The colored regions indicate the different regions used to test N-terminal flexibility.

from the distinct ground states previously identified, or they could result from residual dark-state monomers in the ensembles. The more detailed mathematical analysis of the data given below allowed us to distinguish between these scenarios, and clearly favors the latter.

If the light-activated dimer takes the same form regardless of the initial protein dark state, we can extract its unique scattering profile by minimizing the

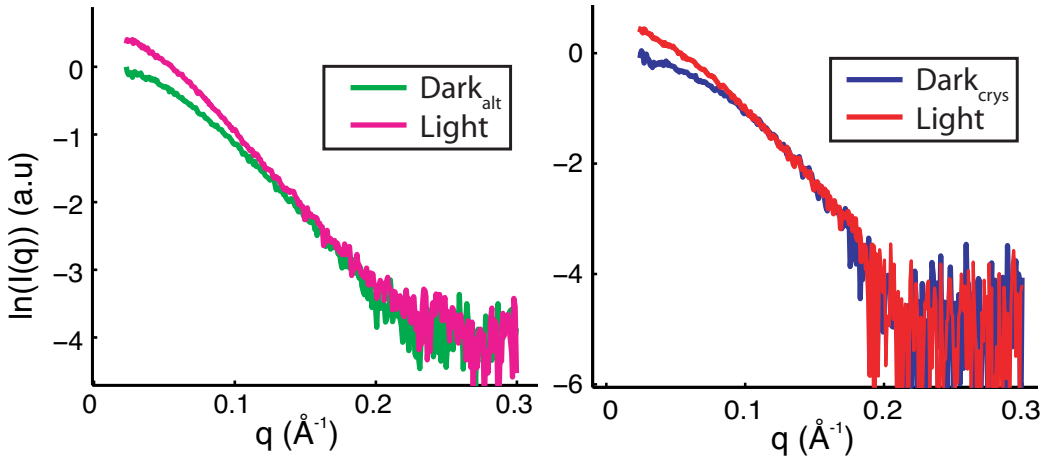

Figure 6.4: Shown at left is the Dark_{alt} state and the associated light-activated data. The $\text{Dark}_{\text{crys}}$ data is shown on the right.

equation

$$\chi_\nu^2 = \sum_q \frac{\left(\left(I_{Lightcrys}(q) - r_1 I_{2PD7}(q)\right) - c\left(I_{Lightalt}(q) - r_2 I_{Darkalt}(q)\right)\right)^2}{\sigma^2_{Lightcrys}(q) + \sigma^2_{Lightalt}(q) + \sigma^2_{Darkalt}(q)} \tag{6.1}$$

Here, $I_x(q)$ are scattering profiles from the different VVD measurements, distinguished by the characteristic scattering of the dark state. For this computation, the scattering profile derived from the 2PD7 structure was employed in place of $\text{Dark}_{\text{crys}}$ to minimize the overall noise. The terms in the denominator represent the errors associated with each measurement and the three parameters r_1, r_2, which represent the fraction of unphotolyzed protein present in each sample and c, which accounts for variations in protein concentration, were varied to minimize χ^2. The values of r_1 and r_2, are 0.62 and 0.83 respectively, indicating that neither initial state significantly inhibits formation of the light-activated state. Thus the scattering profile of the light-excited dimer can be reconstructed by subtracting the dark state contribution from the measured signal using either $I_{Lightcrys}(q) - r_1 I_{2PD7}(q)$ or $c(I_{Lightalt}(q) - r_2 I_{Darkalt}(q))$. The close agreement, seen in Figure 6.5, as well as the χ_ν^2 of 0.68 demonstrates the validity of the original assumption. These four scattering curves can be well described by a model

containing only three states. At this resolution photoactivated VVD forms the same structure regardless of the conformation of the dark state.

The approach also yields a robust scattering profile for the VVD dimer. However, this curve does not match the scattering curve calculated from the crystallized dimer (supplementary information). This may not be surprising, since the light state crystal structure was derived from light excitation of protein crystallized in the dark state. Contacts formed within the dark state crystal may inhibit the motion required to gate dimerization (see discussion). To elucidate these differences, we used established computational methods to reconstruct a representative low resolution structure of the dimers. These methods produced a shape with 2-fold symmetry, shown in Figure 6.6, regardless of whether or not the reconstruction program was seeded with this information. In contrast to the cylindrical crystal structure, this shape has a large bulge along the mirror axis, indicating an expansion of the molecule at the dimer interface. This structure, in combination with the two previously described dark-state ensembles, represents all of the VVD data.

6.2.3 C71V:C183S data

Analysis of the VVD light-state dimer is complicated by a high dissociation constant (2-15 μM) and its involvement in a rapidly dissociating dimer population. The protein is sensitive to oxidation, which can inhibit dimer formation in some cases. To circumvent both of these complications and to better elucidate the light-state dimer structure a C71V:C183S double variant was studied.

C71V:C183S removes surface cysteines which may be involved in oxidative

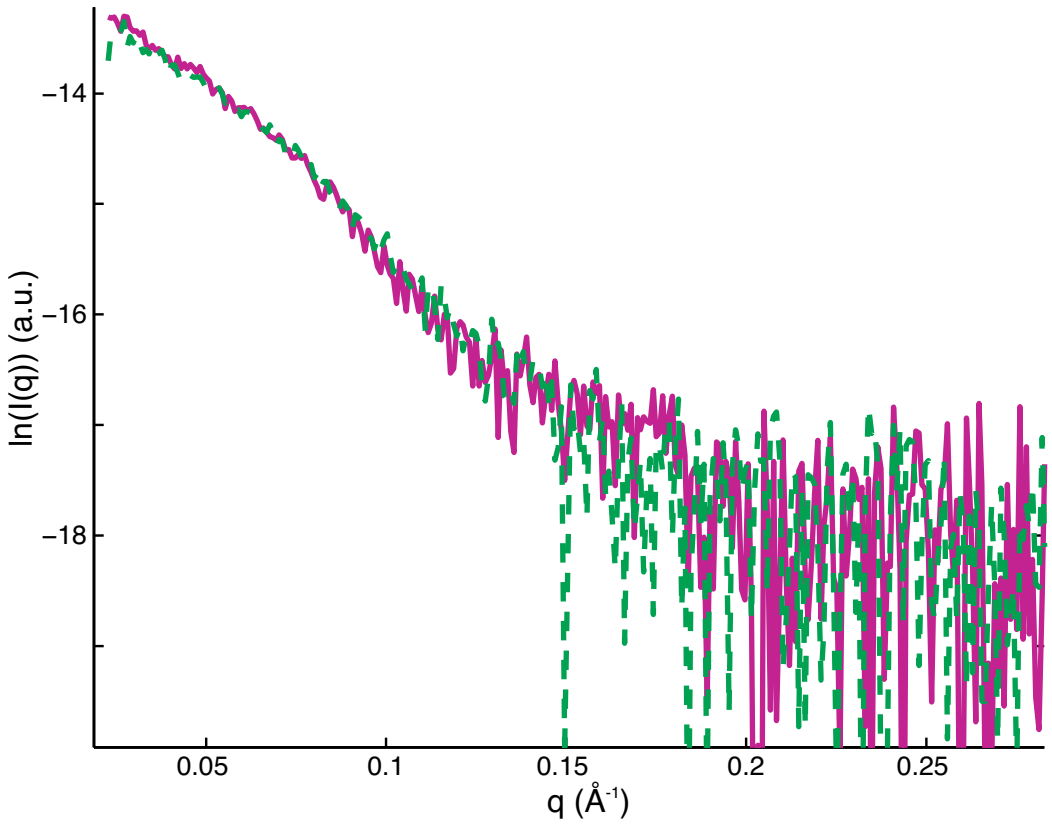

Figure 6.5: Projected scattering from the light-activated dimer, calculated with either the expression $I_{Lightcrys}(q) - r_1 I_{2PD7}(q)$ (solid line) or $c(I_{Lightalt}(q) - r_2 I_{Darkalt}(q))$ (dashed) from the minimization described in the text. The similarity of these states indicates success of the minimization and supports the hypothesis that there is only one light-activated dimer state despite the evidence for two dark state conformation of VVD.

inhibition of the protein. The more significant mutation of Cys71 to Val results in nM dissociation constants. The decrease in the dissociation constant is believed to result from the ability of the Val to occupy positions consistent with both the light- and dark-state conformations of Cys71 observed in crystal structures. Thus, a valine at position 71 predisposes VVD to the light-state conformation even in the absence of light.

To determine the role of the Cys71 conformation in dimer formation and to characterize the light-state dimer, SAXS experiments comparable to those de-

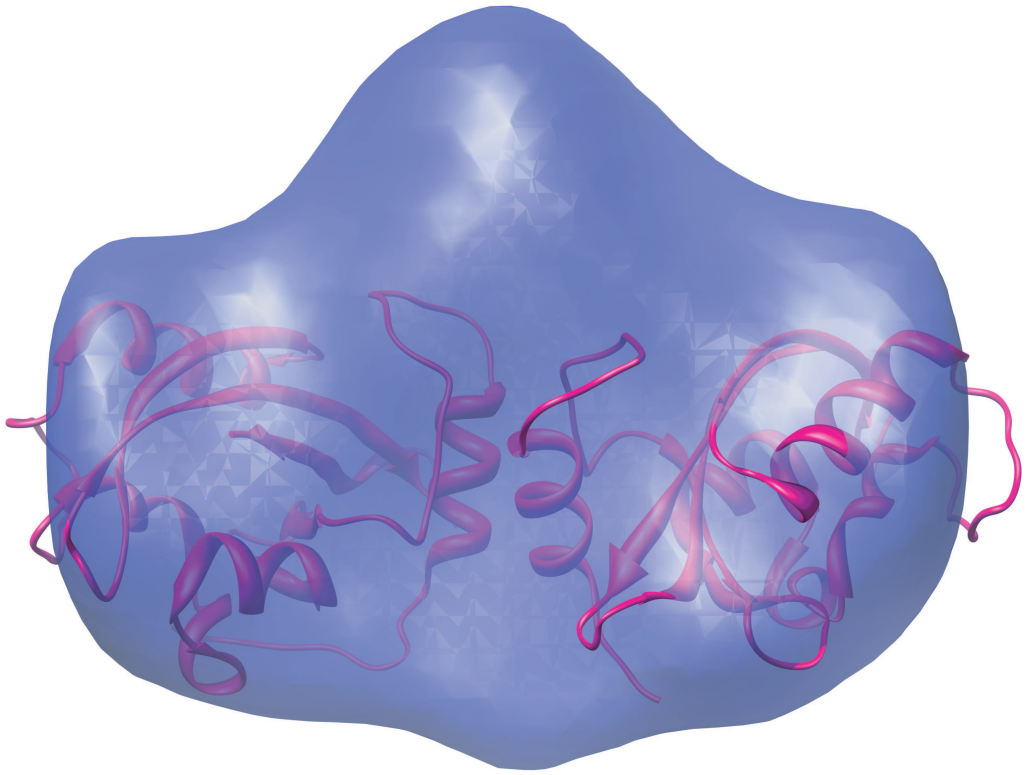

Figure 6.6: Shape envelope of the VVD light-activated dimer reconstructed from the scattering profile extracted by minimization analysis. The mean normalized spatial discrepancy between the ten individual reconstructions should ideally be less than 0.7[16]. This measure of reproduction is 0.55 for this shape envelope, indicating a unique solution to the scattering curve. The crystallized dimer from the light-activated state[9] is superimposed.

scribed above were conducted on both the dark- and light-excited states of the C71V:C183S variant. Although the overall shape of the dark-state scattering profile shown in Figure 6.7 differs from either wildtype curve, the scattering at high q is identical to that of the Dark_{alt} state. Moreover, if the intensity of these high q data is scaled to the Dark_{alt} curve, the low q intensity of the scattering from the variant is higher than that from the wildtype, suggesting a population of dark-state dimers. A two state fit of these data to the Dark_{alt} scattering and the wildtype light-state dimer profile reproduces the data very well (Figure 6.8), unlike a similar fit using the $\text{Dark}_{\text{crys}}$ profile (supplemental data). Thus the

C71V:C183S variant preferentially forms the $Dark_{alt}$ monomer. Furthermore, this conformation is capable of forming a dimer identical at low resolution to the wildtype light-state dimer without excitation. The possibility of accidental light exposure was eliminated upon repetition of the experiment (supplementary data). We deduce that the simultaneous occupation of the light- and dark-state conformations of Cys71 mimics the conformation which gates dimer formation in the wildtype molecule (the light state monomer).

Although dimerization is initiated in the dark-state for the C71S:C183V variant, it is incomplete. Following illumination, $I(0)$ increases by a factor of 1.38 ± 0.03 between the dark and light-activated variant states (Figure 6.8), indicating a larger fraction has dimerized. Previous studies have eliminated the possibility of oligomers of higher order than dimers in VVD and Cys71Val variants. Furthermore, the scattering data are well reproduced by a linear combination of the $Dark_{alt}$ data and the dimer state.

6.2.4 Time-resolved data

VVD wildtype fast flow data

By translating the laser beam relative to the X-ray beam, we acquired time-resolved SAXS data, as described in Materials and Methods. Scattering curves were collected for both variant and wildtype samples. The reducing agent DTT was employed to minimize oxidation, which is believed to inhibit dimerization.

Qualitative inspection of the curves did not reveal obvious time dependent changes in shape, so a quantitative Guinier analysis was used to calculate the

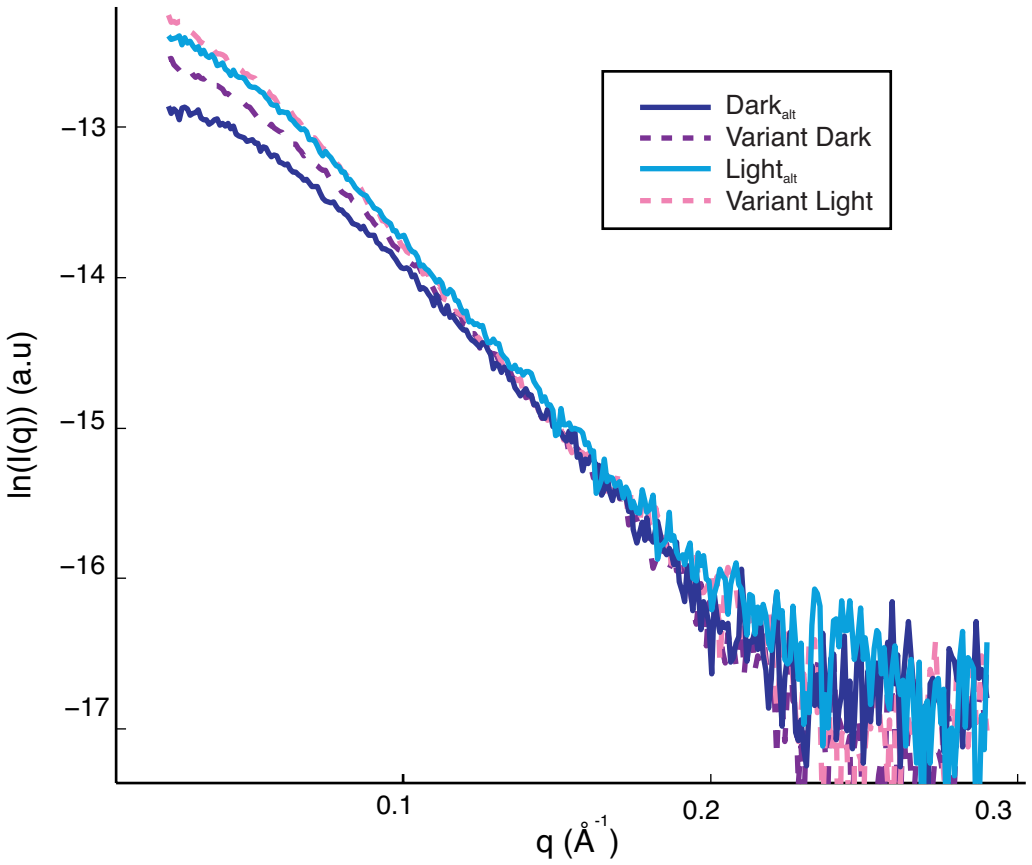

Figure 6.7: SAXS data collected on the VVD variant C71V:C183S compared with the wildtype molecule. Dark colored lines represent scattering from the Dark_{alt} and variant dark-state, while those in lighter shades represent the corresponding light-excited states. Solid lines depict data collected on the wildtype molecule while dashed curves show scattering from the variant.

quantities $I(0)$ and R_g, (see supplementary data for figures). Data were collected as rapidly as 20 ms after excitation for $\text{Dark}_{\text{crys}}$, wildtype VVD and 300 ms after excitation for Dark_{alt} wildtype and the variant. All forms exhibit rapid dimerization, which is nearly complete by the time of the first measurement. Values of $I(0)$ and R_g demonstrate a slight trend towards larger values at the longest times after excitation, consistent with small amounts of additional dimer forming after several seconds. However, the most obvious conclusion from these measurements is that Dark_{alt} and the variant dimerize on a diffusion-limited time scale, as has been reported for the $\text{Dark}_{\text{crys}}$ state[11].

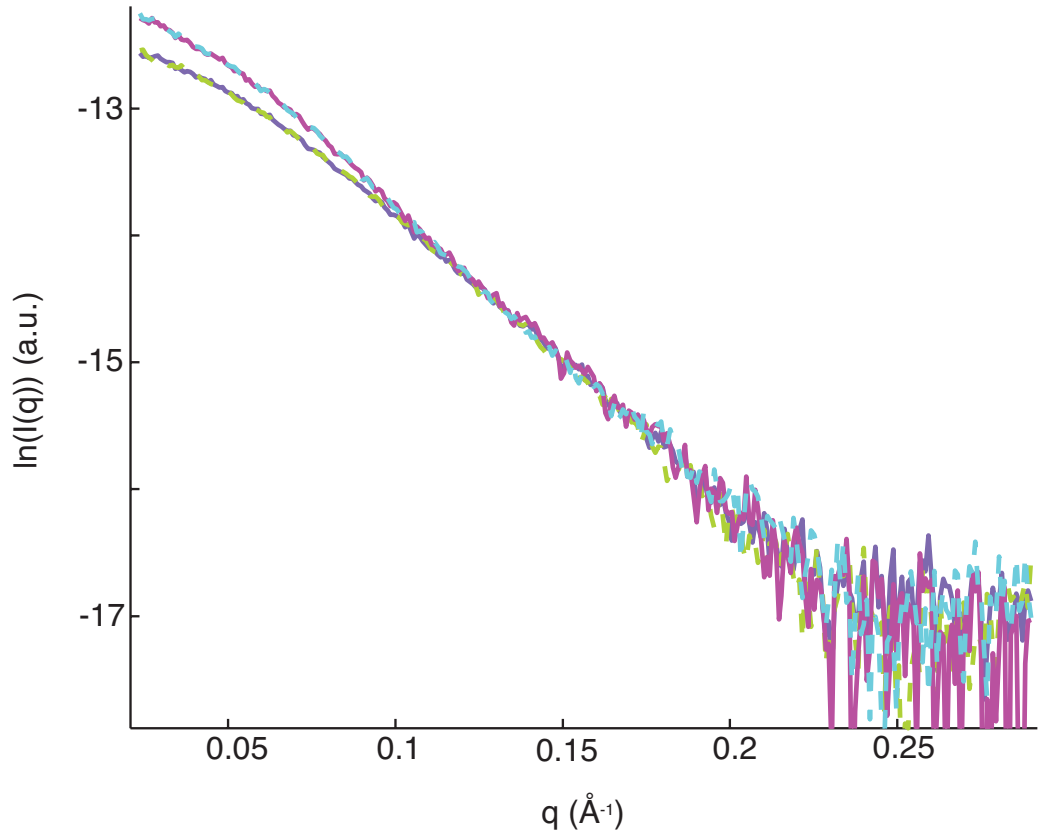

Figure 6.8: Fits to the variant data from a linear combination of scattering from $Dark_{alt}$ and the light-state dimer. In dark solid lines are the data, the lower curve representing the dark state. The light, dashed lines are the fits with χ^2=1.44 and $\chi^2 = 1.91$ for the dark and light states, respectively.

Previous time-resolved SAXS measurements on VVD reflected conformational changes after several seconds, in contrast to current observations[11]. The main difference between these measurements was the inclusion of DTT in the protein buffer. This difference indicates that oxidation is associated with conformational changes in the light excited state.

Sheath flow

In flow cell measurements of this type, it is desirable to prevent sample contact with the walls of the sample cell. The outermost fluid layer moves at very slow

speeds due to the no-slip boundary condition under laminar flow conditions, and causes a small amount of the measured protein ensemble to have a delay time well beyond the sample average. This could cause a "smoothing" in the shape of any feature in the data with a strong time dependence. To mitigate such problems, we built a flow cell which allowed for a sheath of buffer flow to surround the sample, as described in materials and methods. Unfortunately, this protein tends to clump in the narrow sample outlet region of the flow cell, limiting our ability to consistently collect reliable data under these more desirable conditions. However, scattering data useful for comparison was acquired on the variant. Figure 6.9 demonstrates good agreement between short data time points both with and without sheath flow. This confirms that the major conclusion reached from the time-resolved data, namely that dimer formation is largely complete within tens of milliseconds of light excitation, is not being overly influenced by the lack of sheath flow in some measurements.

6.3 Discussion

Time-resolved SAXS profiles of light- and dark-adapted states of VVD-36 and a C71V:C183S SAXS data are effectively described by linear combinations of scattering profiles from just three states: the monomer from the dark-state crystal structure, the same monomer with an extended N-terminus, and a light state dimer with a structure that is more extended than the crystal structure. Taken together, all of these data support a structural model in which the N-terminal α-helix packs weakly against the PAS-scaffold, allowing transient undocking of the helix from the PAS core. We propose that this undocking exposes and stabilizes the dimer interface.

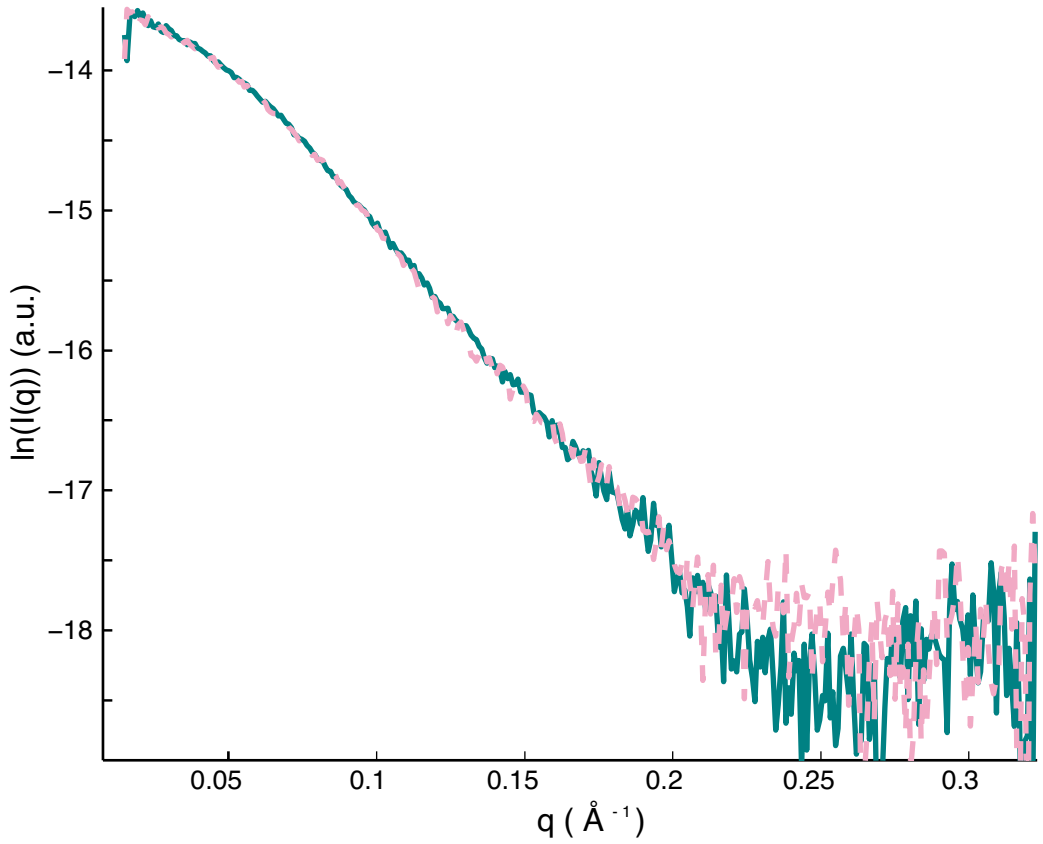

Figure 6.9: Scattering data from the C71V:C173S variant collected with and without a buffer sheath. The dark, solid line represents scattering data collected at t=0.2 s after photoexcitation while using a sheath flow. The light, dashed line is the scattering data collected at t=0.3 s without sheath flow.

First, we discuss the presence of the two distinct dark states in the wildtype population and the valuable clues that their differences provide to the proposed mechanism. While the $\text{Dark}_{\text{crys}}$ scattering profile agrees with the known crystal structure, the Dark_{alt} is well described by molecules where the last 28 N-terminal amino acids are treated as a flexible ensemble. Although modeling at this level can demonstrate agreement between measurement and model, it cannot guarantee that the solution we propose is unique. However, this interpretation is consistent with (1) the structure of the molecule, which shows the N-terminus is near the surface and thus extendable, (2) crystal structures which

suggest alternate conformations of the N-terminus, and (3) similarity to the dimerization of related PAS domains . From these data we propose that motion of these N-terminal residues away from the core gates dimerization. Although we have not yet identified the mechanism for this transformation between the dark states, we are considering several possibilities: oxidation, degradation, and non-equilibrium population of states. Oxidation may influence dimerization and appears to affect structure of the light-excited dimer; thus it is reasonable to suspect it may also impact the dark state molecule. N-terminal degradation can influence the action of the molecule as well[10]. The protein is manufactured in the 'light state'. Return to the 'dark state' is then assessed through spectroscopy which measures the formation of a cysteine photoadduct.. Conformational changes occurring after the adduct breaks, but before the molecule fully returns to its ground state may explain the difference between batches. Despite attempts to probe these phenomena experimentally, we cannot yet state which factor determines the dark-state conformation of wild-type VVD.

More insight into the link between flexibility and dimerization has been gained from scattering studies of a C71V:C183S variant. The scattering profile of the variant dark state is well described by a linear combination of the $Dark_{alt}$ curve and a dimer. Mutation of a cysteine to a valine at a specifically targeted site stabilizes the extended state, facilitating formation of a small fraction of dimer in the dark state, which has the same structure as the light-activated dimer measured in the wildtype. Therefore, the dark-state of the molecule must resemble the molecule that precedes dimer formation, which implies the precursor is an extended form of VVD. Upon light excitation this molecule undergoes a conformational change that further facilitates dimerization. SAXS is capable of measuring only low-resolution structural ensembles; therefore addi-

tional, small changes between $Dark_{alt}$ and the light-state monomer cannot be discerned. Moreover, it is likely that the dark-state monomers ($Dark_{alt}$) that exist in an extended conformation are rapidly sampling a multitude of conformational states. Light-excitation and the Cys71Val variant may stabilize a substate within this molecular ensemble that facilitates dimer formation in light-adapted proteins.

In spite of the distinct states present in dark-acclimated VVD, a common structure for the light-activated dimer is measured. Reconstruction of the dimer curve supports conclusions from previous work that the solution dimer takes a different form from the crystallized dimer. This is not surprising, given that the crystal structure was obtained by light exposure of a crystal containing dark-acclimated VVD. The dimers formed in the dark state crystal may be the result of fortuitous crystal contacts; from our structural models, contact between monomers occurs between the regions that 'gate' dimerization: hence they must be flexible in solution. Dimerization in the dark state locks the gate. At the resolution provided by SAXS we detected a single conformation for the light-activated dimer, implying that formation of the dimer forces the region that is extended in $Dark_{alt}$ but not in $Dark_{crys}$ into a common conformation. The reconstructed envelope depicts a bulge on the mirror axis, which we speculate is due to the associating N-terminal helices of both molecules. The data suggest that this extension precedes dimerization. Such a model is particularly intriguing because of its similarity to the related bacterial LOV domain YtvA[13]. The dimer-interface may involve a β-sheet:β-sheet interaction consistent with other PAS domain dimers (ARNT, HIF, YtvA, etc), with additional contacts made via domain swapping of a α-helix.

Adopting a dimer interface analogous to YtvA would place the H-I loops of the VVD structure in close contact. Such an interface is consistent with solution cross-linking studies, which indicate that introduction of an E171C variant leads to the preferential formation of cross-linked dimers in the presence of light. Previous crystallographic dimers were unable to explain the observed light-dependent formation of disulfide bonds in the E171C variant. Moreover, the molecular envelope suggests that the dimer must be parallel with respect to the orientation of the central β-scaffold. Such an orientation is consistent with YtvA and ARNT homodimers, but in direct opposition to ARNT:HIF heterodimers. Notably, the parallel orientation is consistently present in homodimeric PAS dimers.

VVD has been shown to work in concert with WC1 to regulate circadian rhythms in neurospora crassa. The mechanism for the signaling pathway has yet to be determined. Transient dimerization coupled to conformational dynamics in PAS domains has been observed in phototropins, and HIF:ARNT heterodimers[17]. In phototropins, N-terminal and C-terminal elements akin to the VVD N-terminal α-helix have been shown to undergo large scale conformational changes following photoexcitation. The HIF:ARNT system is intriguing in its similarities. In this case, the HIF:ARNT heterodimer competes with the ARNT homodimer on the signaling pathway. In both cases dimer formation in the full length proteins is believed to be a key event in signal transduction, which might be similar to the way VVD participates in regulation.

The combined data present a model in which photo-excited VVD reorganizes structural elements adjacent to the β-scaffold. The reorganized protein surface is then conducive to formation of a VVD dimer. As an interesting caveat

we have identified two distinct monomer populations within the VVD dark-state. Despite extensive experimentation and data acquired on the two independent species we have not been able to determine a mechanism, which regulates the population of the two states. However, oxidation, N-terminal degradation, or long term effects of light excitation may play a role.

6.4 Materials and Methods

6.4.1 Sample Preparation

Preparation of VVD Variants

The C71V:C183S variant was constructed according to the QuickChange protocal (Stratagene). Resultant mutants were sequenced in their entirety at the Biotechnology Resource Center at Cornell University.

Protein expression and Purification

VVD-36 and C71V:C183S variants were overexpressed in *E. coli* BL21(DE3) cells. Two-liter cultures of the variants were grown to an OD_600 of 0.6-0.8 at 37° C. When the cell density reached 0.6, the cultures were cooled to 18° C and induced with 100 M IPTG. Protein was then expressed for 22 hours prior to harvesting the cells.

24-liters of both VVD-36 and C71V:C183S were prepared via the above protocal and lysed in buffer containing 13% glycerol, 300 mM NaCl, 50 mM Hepes

pH 8.0 and 5 mM imidazole pH 8, as the sonicated and soluble cell lysate was factionated by centrifugation. The supernatant was then collected and protein purified via Ni:NTA affinity chromatography. eluted VVD was subsequently treated with 1 unit of thrombin/mg of protein for 6 hours in buffer containing 2 mM DTT, 13% glycerol, 150 mM NaCl, 50 mM Hepes pH 8.0 and 100 mM imidazole. The protein was then purified on a Superdex 75 26/60 Hi-load column and concentrated to 5 mg/mL. Final protein samples contained 5 mM DTT.

6.4.2 Time-resolved SAXS measurements

Small angle X-ray scattering data was collected at the G1 beamline at the Cornell High Energy Synchrotron Source (CHESS) at an energy of 8 keV. A continuous flow cell made of a 1 mm polyester tube (Advanced Polymers, Inc., VT)[18] was employed to collect time resolved data, using a method described in Lamb et al. [11] . A 473 nm laser from Holograms and Laser, International (Houston, TX) was aligned with the X-ray beam using a 20 μm slit to confirm their coincidence. The X-ray beam was focused with a glass capillary[19;20] for better position definition. The laser focal point was then moved against the direction of fluid flow to introduce delay. Eight 30 s X-ray exposures were collected on the protein for each sample, to improve signal to noise and ensure reproducibility. A pin diode was mounted onto the X-ray beamstop to measure changes in beam intensity.

6.4.3 Data analysis

Images were converted to scattering profiles of intensity (I) as a function of q, where $q = \frac{4\pi \sin(2\theta)}{\lambda}$, with θ being half the scattering angle and λ is the X-ray wavelength. An image of a silver stearate[21] scattering ring collected under the same beam conditions was used to find the beam center and determine the radial calibration. Each scattering profile was normalized for changes in beam intensity and checked for reproducibility. Final scattering profiles were obtained by first averaging images, then converting the data to intensity versus q, and subtracting the buffer background. This analysis was carried out using MATLAB (The Mathworks, Natick, MA).

Using the pin diode to normalize the data allows meaningful comparison of $I(0)$ from data collected within a period of a few hours. However, over the course of many hours or days this becomes less valid. In general, when comparing different scattering data, no scaling additional to the PIN diode was performed for data taken on the same sample on the same day.

When the data is plotted as a Guinier plot, $\log(I(q))$ versus q^2, the plot is approximately linear at low angles[22]. Typically, this approximation is considered valid for $qR_g < 1.3$. Fitting to this line allows extrapolation to $I(0)$ and provides the slope, which is proportional to the radius of gyration R_g. Errors on these quantities were determined by propagating the 95% confidence intervals from the slope and y-intercept of the line fit with MATLAB.

GNOM, DAMMIN, DAMAVER, and EOM are all analysis tools for scattering data made available by the Biological Small Angle Scattering group at the European Molecular Biology Laboratory[16](`http://www.embl-hamburg.`

de/ExternalInfo/Research/Sax/software.html). Conversion of scattering data to P(r) was done with GNOM[23]. A GNOM user must provide a maximum particle dimension (D_{max}) as an input. In general D_{max} was found by testing values based on the R_g determined from Guinier analysis and choosing the one which best maximized the default regularization parameters used by GNOM. We determined D_{max} to within 5 Å with this method. GNOM was also a precursor to reconstructing the dimer shape envelope, and to find statistical errors on the scattering data points based on the dispersion of points.

All minimizations described in the text were performed using the fmin function in MATLAB. Statistical errors for the analysis were determined using GNOM based on the noise in the data. Degrees of freedom, ν were calculated by subtracting the number of fitting parameters from the number of points in a single scattering curve. Reconstructions were carried out by running DAMMIN[24] ten times on the output from GNOM, and then using DAMAVER to average the results and Situs[25] to generate a shape envelope. Analysis with EOM was carried out using the 3D72 crystal structure. This was used in place of the 2PD7 structure because multiple electron densities of the C108 made the PDB difficult for EOM to handle. Scattering from the two structures is indistinguishable.

REFERENCES

1. Lamb, J. S., B. D. Zoltowski, S. A. Pabit, L. Li, B. R. Crane, and L. Pollack. "Illuminating conformational changes of a LOV domain protein with photocoupled small angle X-ray scattering." in preparation.

2. Taylor, B. and I. Zhulin. "PAS domains: internal sensors of oxygen, redox potential, and light." *Microbiol. Mol. Biol. R.*, **63**(2): 479, 1999.

3. Cheng, P., Q. He, Y. Yang, L. Wang, and Y. Liu. "Functional conservation of light, oxygen, or voltage domains in light sensing." *Proc. Natl. Acad. Sci. U. S. A.*, **100**(10): 5938–43, 2003.

4. Huala, E. "*Arabidopsis* NPH1: A protein kinase with a putative redox-sensing domain." *Science*, **278**(5346): 2120–2123, 1997.

5. Crosson, S. and K. Moffat. "Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction." *Proc. Natl. Acad. Sci. U.S.A.*, **98**(6): 2995–3000, 2001.

6. Loros, J. J. and J. C. Dunlap. "Genetic and molecular analysis of circadian rhythms in Neurospora." *Annu. Rev. Physiol.*, **63**: 757–94, 2001.

7. Avila-Perez, M. "Blue light activates the B-dependent stress response of *Bacillus subtilis* via YtvA." *J. Bacteriol.*, **188**(17): 6411–6414, 2006.

8. Heintzen, C., J. J. Loros, and J. C. Dunlap. "The PAS protein VIVID defines a clock-associated feedback loop that represses light input, modulates gating, and regulates clock resetting." *Cell*, **104**(3): 453–464, 2001.

9. Zoltowski, B. D., C. Schwerdtfeger, J. Widom, J. J. Loros, A. M. Bilwes, J. C. Dunlap, and B. R. Crane. "Conformational switching in the fungal light sensor Vivid." *Science*, **316**(5827): 1054–7, 2007.

10. Zoltowski, B. and B. Crane. "Light activation of the LOV protein vivid generates a rapidly exchanging dimer." *Biochemistry*, **47**(27): 7012–7019, 2008.

11. Lamb, J., B. D. Zoltowski, S. A. Pabit, B. R. Crane, and L. Pollack. "Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering." *J. Am. Chem. Soc.*, 2008. Doi:10.1021/ja804236f.

12. Ma, X., N. Sayed, P. Baskaran, A. Beuve, and F. V. D. Akker. "PAS-mediated dimerization of soluble guanylyl cyclase revealed by signal transduction histidine kinase domain crystal structure." *Journal of Biological Chemistry*, **283**(2): 1167–1178, 2007.

13. Möglich, A. and K. Moffat. "Structural Basis for Light-dependent Signaling in the Dimeric LOV Domain of the Photosensor YtvA." *J. Mol. Biol.*, **373**(1): 112–126, 2007.

14. Strickland, D., K. Moffat, and T. R. Sosnick. "Light-activated DNA binding in a designed allosteric protein." *Proc. Natl. Acad. Sci. U. S. A.*, **105**(31): 10709–14, 2008.

15. Nakasako, M., T. Iwata, D. Matsuoka, and S. Tokutomi. "Light-induced structural changes of LOV domain-containing polypeptides from Arabidopsis phototropin 1 and 2 studied by small-angle X-ray scattering." *Biochemistry*, **43**(47): 14881–90, 2004.

16. Konarev, P. V., M. V. Petoukhov, V. V. Volkov, and D. I. Svergun. "ATSAS 2.1, a program package for small-angle scattering data analysis." *J. Appl. Cryst.*, **39**: 277–286, 2006.

17. Card, P. B., P. J. A. Erbel, and K. H. Gardner. "Structural basis of ARNT PAS-B dimerization: use of a common beta-sheet interface for hetero- and homodimerization." *Journal of Molecular Biology*, **353**(3): 664–77, 2005.

18. Kalinin, Y., J. Kmetko, A. Bartnik, A. Stewart, R. Gillilan, E. Lobkovsky, and R. Thorne. "A new sample mounting technique for room-temperature macromolecular crystallography." *J. Appl. Cryst.*, **38**(2): 333–339, 2005.

19. Engstrom, P., S. Larsson, A. Rindby, A. Buttkewitz, S. Garbe, G. Gaul, A. Knochel, and F. Lechtenberg. "A submicron synchrotron x-ray beam generated by capillary optics." *Nucl. Instrum. Methods Phys. Res., Sect. A*, **302**: 547–552, 1991.

20. Lamb, J. S., S. Cornaby, K. Andresen, L. Kwok, H. Park, X. Qiu, D. M. Smilgies, D. H. Bilderback, and L. Pollack. "Focusing capillary optics for use in solution small-angle X-ray scattering." *J. Appl. Cryst.*, **40**: 193–195, 2007.

21. Vand, V., A. Aitken, and R. K. Campbell. "Crystal structure of silver salts of fatty acids." *Acta Cryst.*, **2**: 398–403, 1949.

22. Guinier, A. and G. Fournet. *Small-Angle Scattering of X-Rays*. John Wiley and Sons., Inc., New York, 1955.

23. Svergun, D. I. "Determination of the regularization parameter in indirect-transform methods using perceptual criteria." *J. Appl. Cryst.*, **25**: 495–503, 1992.

24. Svergun, D. "Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing." *Biophys. J.*, **76**(6): 2879–2886, 1999.

25. Wriggers, W. and P. Chacón. "Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering." *J. Appl. Cryst.*, **34**: 773–776, 2001.

CHAPTER 7

FUTURE WORK AND OTHER APPLICATIONS

This document describes experiments conducted on many molecules using a variety of time-resolved SAXS methods. The results presented provide insight into the behavior of VVD and other biomolecules. However, as is always the case in scientific investigation, the data collected lead to more questions which require answers. Also, the unique methods used to conduct these experiments could be applied to other interesting molecules.

7.1 Small Angle X-ray Scattering Measurements on Viral Fusion Proteins

7.1.1 Importance of Fusion proteins

Enveloped viruses invade target cells by actively merging the viral and cell membranes together. This function is accomplished by viral fusion proteins. These surface proteins are typically triggered by either binding of a cell surface receptor or a drop in pH[1]. The activated proteins then undergo significant structural rearrangements that are associated with membrane fusion[2]. Influenza hemagglutinin (HA) is a very well studied example of a class I viral fusion protein. Influenza is brought to the endosome upon cell binding where HA transforms in response to acidification of the compartment, causing membrane fusion and infection. Because these proteins are found on the viral surface they are antibody targets, and because they perform a crucial function they are

therapeutic targets as well[3]. These proteins are also interesting for engineering drug delivery systems, since they naturally function to deliver the contents of a virus to a cell. Thus viral fusion proteins are molecules which attract a good deal of scientific interest.

Viral fusion proteins are ideal candidates for study by the rapid mixing techniques discussed in chapter 2. The pH activated proteins represent biologically relevant systems which respond *in vivo* to the exact stimulus which rapid mixing can provide. Because these proteins have membrane-bound components, many are difficult to crystallize fully. With the aid of detergents[4;5] SAXS could offer low-resolution structures which can complement the high-resolution parts solved with crystallography. The size of the conformational change needed to achieve membrane fusion should be large enough to be readily measured with SAXS.

7.1.2 SARS spike protein

To test the feasibility of using SAXS to measure the conformational changes of viral fusion proteins, we collected static SAXS data on the spike protein from the SARS coronavirus at pH 7 and pH 5. While the *in vivo* role of pH in the activation of this fusion protein is unclear[6], the ectodomain is known to irreversibly form a trimer in response to a drop in pH[7]. SAXS measurements were collected to observe this structural change.

We collected data with 10 keV x-rays at the G1 beamline at CHESS. Samples were contained in 2 mm cuvettes with silicon nitride windows. Four 5 s exposures were collected on each sample and the scattering intensities were checked

for consistency before averaging. Samples were prepared at approximately 7 mg/mL and diluted by factors of 2 and 4 to check for concentration dependence. The detergent-like sample solution was difficult to load into sample cells without dilution, so qualitative shape comparisons were performed to check for signs of aggregation or interparticle interference.

The resulting SAXS data are shown in Figure 7.1. Differences are easily perceived between the pH 7 and pH 5 samples. This difference is reflected quantitatively by the 11 Å increase in R_g obtained by a Guinier fit to the data. Despite an upturn in the low q data which can indicate aggregation, no concentration dependence was observed. The low q feature could be caused by background scatter from the nitride window or parasitic scatter from beam-defining slits. This region was not included when fitting the data.

Reconstructions of the data were carried out using ATSAS[8] as described in Chapters 6 and 4. When reconstructing the pH 5 data it was necessary to specify three-fold symmetry and a prolate particle shape, both of which are available options with DAMMIN. Without the latter option, half of the reconstruction attempts found a very different, oblate shape. The results are shown in Figure 7.2 compared with electron microscopy (EM) images of the same molecules[7]. The MNSD for each averaged reconstruction is 0.66 for the pH 7 computation and 0.73 for pH 5. These low values demonstrate that the results are unique[9]. The EM images show a number of similar conformations of the spike protein. The SAXS reconstructions show only an average for the full sample. Once this is taken into account, the reconstructions and EM images seem to agree in shape, though the size of the SAXS reconstructions is a bit bigger. Some of this size increase may be due to a hydration shell around the molecule in aqueous solu-

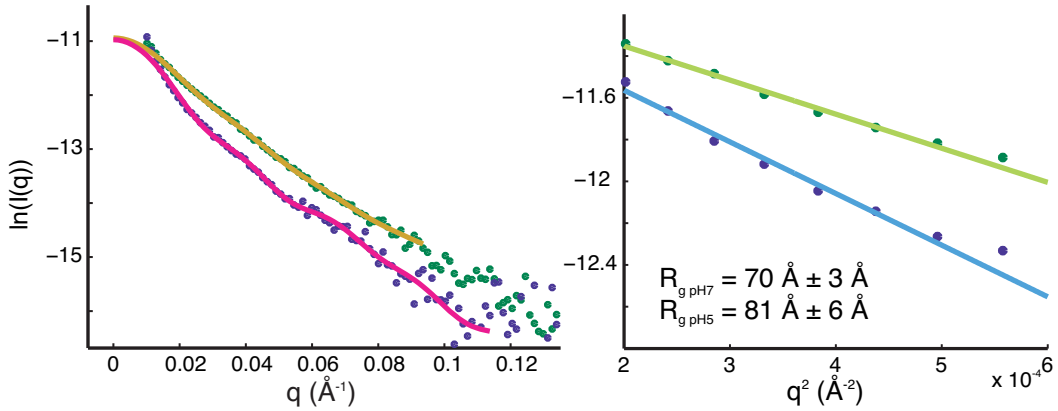

Figure 7.1: SAXS data collected on the SARS spike protein at pH 7 and pH 5. On the left the data are plotted with the smoothed GNOM fits to the data, which are used in the reconstruction process. On the right are the Guinier fits with the data. Dots represent data points while solid lines are fits. Green dots show pH 7 data; blue dots are pH 5.

tion.

7.1.3 Time-resolved SAXS data on influenza hemagglutinin

Influenza HA is an extensively studied fusion protein. Proper HA function is crucial for influenza to successfully infect a host. It is found in different forms in each strain of the virus. The structure of the ectodomain at neutral pH was solved by x-ray crystallography in 1981 by Wilson et al. [10]. Despite widespread interest, the post-fusion structure of the protein remains unknown. Crystal structures of pieces of HA at pH 5 [11;12], electron micrographs [12;13], and a variety of other studies have led to a model of pH mediated fusion whereby the N-terminus of each HA_2 subdomain moves from one end of the protein to another upon acidification. This conformational change of the ectodomain is ideal for studying with SAXS using one of the time-resolved pH jump methods described in this document. In addition to providing information on possible

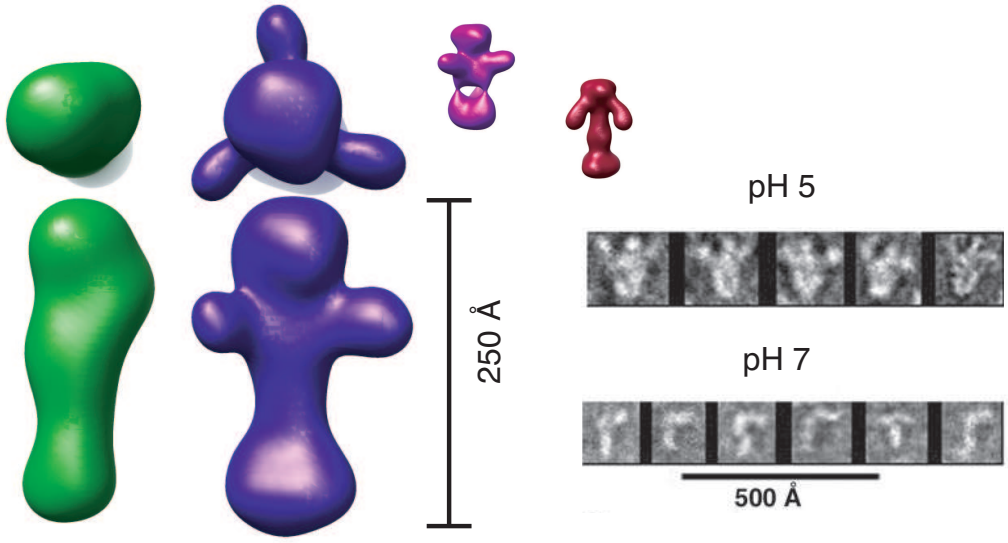

Figure 7.2: Reconstructions of the SARS spike protein. In green are two views of the averaged reconstruction of the pH 7 protein; in blue are depictions of the pH 5 molecule. In red and magenta are two individual reconstructions of the pH 5 data which vary somewhat from the average. On the right are EM images published by Li et al. [7] for comparison.

fusion intermediates, a pH jump may capture data on the low-pH conformation before it precipitates.

We performed preliminary experiments using a stopped flow mixer setup described in Appendix A at the CHESS G1 beamline. We collected data on the ectodomain of HA at a concentration of approximately 3 mg/mL after a pH jump from pH 7 to pH 5. Two 100 ms exposures were collected for each data point to conserve sample and check reproducibility. The results are shown in Figure 7.3. At this initial concentration signal from HA was negligible by 500 ms after mixing, indicating the expected post-acidification precipitation was complete. However, before this a change between the shape of the SAXS curve collected at neutral pH and 200 ms after acidification was observed. Although the curve was originally very noisy, making exact interpretation of this difference unclear, it demonstrates that we can capture a conformation change associated

with the fusion transition of HA using SAXS.

7.1.4 Acknowledgments

Samples of viral fusion proteins were provided by Professor Whittaker in the School of Veterinary Medicine, Cornell. In particular, Victor Chu, Shoshannah Roth, and Xianjie Sun prepared the protein. Arthur Woll, the beamline scientist at G1 helped set up the experiment.

7.2 Potential applications of methods

7.2.1 Concentric flow cell with laser-induced excitation

In Chapters 2, 5 and 6 we describe a concentric flow cell which uses light excitation to initiate molecular changes in VVD which are then measured by SAXS. Such a method can be expanded to a wide variety of other molecules. There are a wide variety of light activated proteins which might be studied with such a system. In addition to the many blue-light activated PAS-LOV proteins, there are rhodopsins, phycobiliproteins, and photoactive yellow protein (PYP), to name a few. Such proteins perform a wide variety of functions including photosynthesis, photoprotection and photosensing[14]. Any light activated protein with a large conformational change could be studied using the flow-cell method with SAXS.

Also, a continuous flow cell as described above could be employed for use

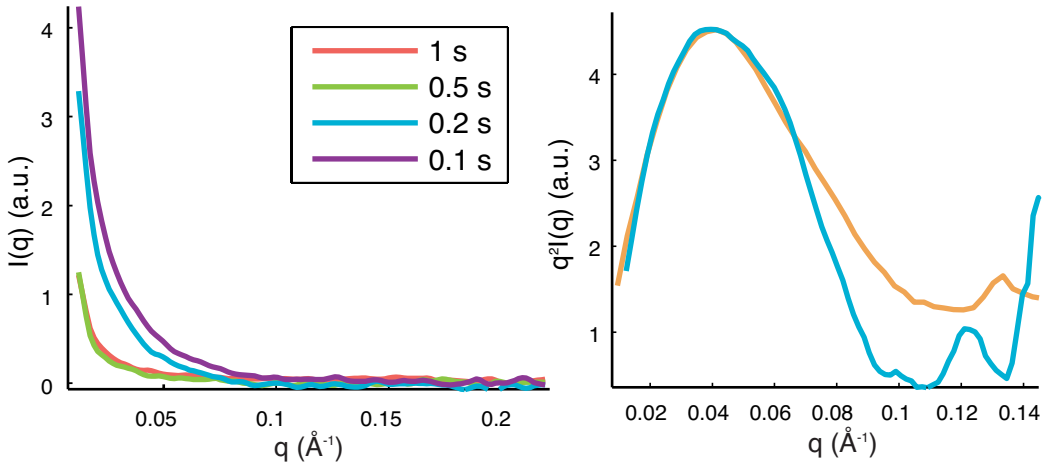

Figure 7.3: Results for preliminary time-resolved HA experiments. The figure on the left shows SAXS data collected at several times after acidification, with intensity decreasing with increasing time. On the right, the shape of the scattering curve collected 200 ms after acidification (blue) is compared to pH 7 data (orange).

with optical techniques such as spectroscopy. This has already been described in Toepke et al. [15], but its utility in investigating light-active proteins is particularly worth noting, because the spectroscopic changes they undergo in response to light are directly related to protein function. Many spectroscopic experiments rely on repetitions of a measurement on a protein with a fast photocycle to build signal-to-noise. For example, Kottke et al. [16] needed to employ a mutant to study the photocycle of a Phot1 receptor from *C. reinhardtii* due to the long photocycle of the wildtype. A continuous flow cell such as we have described avoids these problems.

A photocoupled flow cell could also be used to conduct experiments with molecules that do not respond directly to light through the use of caged compounds. These chemicals are typically photolyzed by UV light to release relevant compounds such as ions, NTPs, neurotransmitters, and acids[17]. Many biomolocules react to such triggers, and could be dissolved with the caged pre-

cursors so a reaction could be triggered and studied.

7.2.2 Analysis of time-resolved data with reconstructions

Chapter 4 demonstrates the reconstruction of time-resolved data into low-resolution shape envelopes. Reconstructions are becoming widely used, and this demonstrates another potential application. While other methods such as R_g or SVD component analysis might be useful to map the time-resolved changes, reconstructions could be very useful in projecting intermediate states discovered by time-resolved methods. They could also be very useful in presenting the mean shape of low-order ensembles where individual components are too varied to extract by methods like SVD.

7.3 Future science goals and open questions

7.3.1 Viral fusion proteins

This chapter describes preliminary experiments on viral fusion proteins. Many of these proteins are ideally studied with a pH jump method because their response to pH is extremely biologically relevant. While the connection between acidification, conformational change, and cell invasion is well established, the exact nature of the structural change remains unclear. Electron microscopy has been employed to look at the pH 5 state, which has been unavailable to crystallography[18]. Our experiments show that SAXS could be employed for further study of the low-pH conformation, and to elucidate any intermediates. While

the stopped-flow method only produced low signal-to-noise preliminary data, an analogous experiment with a rapid diffusive mixer[19] might add to the results. These techniques could then be used on less studied proteins, such as the vesicular stomatitis virus glycoprotein (VSV G).

7.3.2 Vivid

Our experiments on VVD demonstrated the existence of two dark states, without revealing an obvious cause of the difference. This open question could be investigated with static SAXS or alternate techniques which are sensitive to the size and extension of the molecule. One variable that could be explored is oxidation. LOV domain proteins, including VVD, are by definition sensitive to oxidation[20;21]. While the absolute extent of protein oxidation is difficult to control, an experiment could be designed to look for changes in molecular conformation due to relative differences in oxidation levels.

Further time-resolved SAXS experiments could be conducted on VVD as well. UV light is supposed to induce a back-reaction from light-state to dark[22]. Such a process studied by SAXS could provide insights into conformational changes in the back-reaction that follow breaking the photoadduct. Also, a faster time-resolved study could elucidate the monomer which precedes dimerization. Such a study might require reducing sample consumption of the overall flow cell.

REFERENCES

1. Earp, L., S. Delos, H. Park, and J. White. "The many mechanisms of viral membrane fusion proteins." *Curr. Top. Microbiol. Immunol*, **285**: 25–66, 2005.

2. Colman, P. and M. Lawrence. "The structural biology of type I viral membrane fusion." *Nature Reviews Molecular Cell Biology*, 2003.

3. Eckert, D. and P. Kim. "Mechanisms of viral membrane fusion and its inhibition." *Ann. Rev. Biochem.*, **70**: 777–810, 2001.

4. Bu, Z. and D. Engelman. "A method for determining transmembrane helix association and orientation in detergent micelles using small angle x-ray scattering." *Biophys. J.*, **77**(2): 1064–1073, 1999.

5. Lipfert, J., L. Columbus, V. B. Chu, and S. Doniach. "Analysis of small-angle X-ray scattering data of protein-detergent complexes by singular value decomposition." *J. Appl. Cryst.*, **40**: S235–S239, 2007.

6. Simmons, G., J. D. Reeves, A. J. Rennekamp, S. M. Amberg, A. J. Piefer, and P. Bates. "Characterization of severe acute respiratory syndrome-associated coronavirus (SARS-CoV) spike glycoprotein-mediated viral entry." *Proc. Natl. Acad. Sci. USA*, **101**(12): 4240–4245, 2004.

7. Li, F., M. Berardi, W. Li, M. Farzan, and P. Dormitzer. "Conformational states of the severe acute respiratory syndrome coronavirus spike protein ectodomain." *J. Virol.*, **80**(14): 6794–6800, 2006.

8. Konarev, P. V., M. V. Petoukhov, V. V. Volkov, and D. I. Svergun. "ATSAS 2.1, a program package for small-angle scattering data analysis." *J. Appl. Cryst.*, **39**: 277–286, 2006.

9. Volkov, V. V. and D. I. Svergun. "Uniqueness of ab initio shape determination in small-angle scattering." *J. Appl. Cryst.*, **36**: 860–864, 2003.

10. Wilson, I. A., J. J. Skehel, and D. C. Wiley. "Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution." *Nature*, **289**(5796): 366–73, 1981.

11. Bullough, P., F. Hughson, J. Skehel, and D. Wiley. "Structure of influenza haemagglutinin at the pH of membrane fusion." *Nature*, **371**(6492): 37–43, 1994.

12. Chen, J., J. J. Skehel, and D. C. Wiley. "N- and C-terminal residues combine in the fusion-pH influenza hemagglutinin HA(2) subunit to form an N cap that terminates the triple-stranded coiled coil." *Proc. Natl. Acad. Sci. USA*, **96**(16): 8967–72, 1999.

13. Skehel, J. J., P. M. Bayley, E. B. Brown, S. R. Martin, M. D. Waterfield, J. M. White, I. A. Wilson, and D. C. Wiley. "Changes in the conformation of influenza virus hemagglutinin at the pH optimum of virus-mediated membrane fusion." *Proc. Natl. Acad. Sci. USA*, **79**(4): 968–72, 1982.

14. Hellingwerf, K., W. Hoff, and W. Crielaard. "Photobiology of microorganisms: how photosensors catch a photon to initialize signalling." *Mol. Microbiol.*, **21**(4): 683–693, 1996.

15. Toepke, M. W., S. H. Brewer, D. M. Vu, K. D. Rector, J. E. Morgan, R. B.

Gennis, P. J. A. Kenis, and R. B. Dyer. "Microfluidic flow-flash: Method for investigating protein dynamics." *Anal. Chem.*, **79**(1): 122–128, 2007.

16. Kottke, T., J. Heberle, D. Hehn, B. Dick, and P. Hegemann. "Phot-LOV1: photocycle of a blue-light receptor domain from the green alga Chlamydomonas reinhardtii." *Biophys. J.*, **84**(2 Pt 1): 1192–201, 2003.

17. Ellis-Davies, G. "Caged compounds: photorelease technology for control of cellular chemistry and physiology." *Nat. Methods*, 2007.

18. Böttcher, C., K. Ludwig, A. Herrmann, and M. van Heel. "Structure of influenza haemagglutinin at neutral and at fusogenic pH by electron cryomicroscopy." *FEBS Lett.*, 1999.

19. Pollack, L., M. W. Tate, N. C. Darnton, J. B. Knight, S. M. Gruner, W. A. Eaton, and R. A. Austin. "Compactness of the denatured state of a fast-folding protein measured by submillisecond small-angle X-ray scattering." *Proc. Natl. Acad. Sci. U. S. A.*, **96**: 10115–10117, 1999.

20. Huala, E. "*Arabidopsis* NPH1: A protein kinase with a putative redox-sensing domain." *Science*, **278**(5346): 2120–2123, 1997.

21. Zoltowski, B. and B. Crane. "Light activation of the LOV protein vivid generates a rapidly exchanging dimer." *Biochemistry*, **47**(27): 7012–7019, 2008.

22. Zoltowski, B. D., C. Schwerdtfeger, J. Widom, J. J. Loros, A. M. Bilwes, J. C. Dunlap, and B. R. Crane. "Conformational switching in the fungal light sensor Vivid." *Science*, **316**(5827): 1054–7, 2007.

CHAPTER 8

CONCLUSION

We have applied time-resolved small angle x-ray scattering to a number of interesting problems. A coaxial flow cell was used to couple SAXS with light excitation. Previously, most time-resolved SAXS methods have employed some form of rapid mixing to induce sample changes. This technique is extremely valuable for the many molecules that respond to changes in pH, salt levels, or in response to small ligands. However, this optically transparent device allows the addition of the many interesting light-sensitive molecules to the body of potential applications for time-resolved SAXS methods.

We have also shown that even low-resolution time-resolved data can be successfully reconstructed. Reconstructions have many potential applications, some of which are described in this document. For example, it is often speculated that the conformational change of HA must involve an intermediate. Using time-resolved SAXS, scattering for this intermediate might be collected and a low resolution shape could then be reconstructed. This could help place the fragments which have been crystallized into an overall picture of pH-dependent fusion.

Using our photocoupled flow cell with SAXS, we have made significant contributions towards understanding the blue-light photoreceptor VVD. This protein is responsible for circadian clock and carotenoid pigment production in the fungus *neurospora crassa*[1]. The role of the protein in the circadian clock and carotenoid pigment production is known through mutation experiments on the fungus, but the mechanism for protein action has yet to be elucidated. By demonstrating the rapid formation of a dimer upon photoexcitation, we pro-

vide evidence that signaling by this PAS-LOV domain is facilitated through self-interaction or possibly interaction with other molecules upon light excitation. Such interactions have been proposed as a general mechanism for PAS-LOV domain signaling[2].

Two dark states, one of which is associated with the light state, contributes to a model of the structural rearrangements which lead to the VVD dimer, a likely factor in its signaling pathway. In this model, excitation at the photocenter of the molecule is propagated to the N-terminus of the protein causing a significant conformational change. The PAS protein family is a large collection of sensory proteins found in all kingdoms of life. The structural homology between PAS domains is high, implying the domain form is useful in transducing signals between a sensory domain and a mechanistic one, regardless of the stimulus or the output[2]. Information about the structure and function of an individual PAS domain is relevant to this protein family.

This work describes the overall structural changes in the protein VVD. It also raises future questions about the role of different structures in VVD, and outlines a new means by which to explore light-excited proteins. Various time-resolved SAXS methods and analysis of the data have been used to probe the dynamics and structure of biologically relevant molecules. These techniques can be used to approach diverse systems beyond PAS domain proteins and study large-scale molecular dynamics. We have provided new insight into the workings a protein that are relevant to of a very important family of molecules and discussed experimental and analytical techniques which can further probe a wide variety of systems.

REFERENCES

1. Loros, J. J. and J. C. Dunlap. "Genetic and molecular analysis of circadian rhythms in Neurospora." *Annu. Rev. Physiol.*, **63**: 757–94, 2001.

2. Taylor, B. and I. Zhulin. "PAS domains: internal sensors of oxygen, redox potential, and light." *Microbiol. Mol. Biol. R.*, **63**(2): 479, 1999.

CHAPTER 9

SUPPORTING FIGURES AND TABLES

The following figures and tables were included as supporting materials for the papers presented in this dissertation

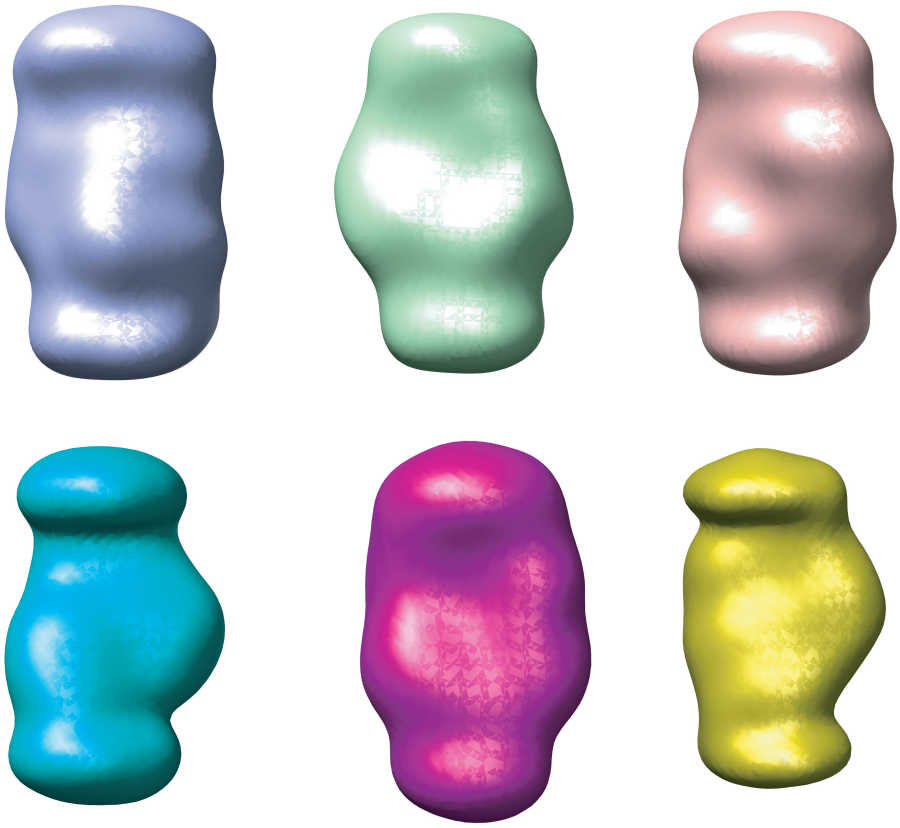

Figure 9.1: Shown are reconstructions of the 92, 125, and 168 ms scattering data from left to right. Those in the top row exclude data from q $>$ 0.18 Å^{-1}. The MNSDs for these averages are 0.589, 0.547, 0.599 respectively[1].

Table 9.1: [a] Normalized to dark state value for ease of comparison. [b] Inactive control mutant. [c] The percent monomer to dimer was projected from the quotient of $I(0)$ in the light and dark states. This estimate is based on the assumption that only monomers and dimers are present in the solution and the conversion accounts for the entirety of the change. Data sets 1 and 2 were taken with different batches of protein during different beamtimes. Guinier analysis was conducted using MATLAB by fitting a line to ln(I) as a function of q^2 at the lowest scattering angles where a line was found to best represent the data. To ensure the results were comparable to each other, the same range was fit for each curve, from q=0.050 to q=0.078. The error bars are propagated from the 95% confidence bounds generated by MATLAB from the statistics of the fit. Interpretation of the radius of gyration is ambiguous due to the mixed nature of the light state, particularly given the large uncertainty in the proportion of monomers to dimers[2].

	Time after excitation	$I(0)$ (a.u.)[a]	% monomer → dimer[c]
Data WT (slow) Set 1	Dark state 1	1.0 ± 0.2	–
	Dark state 2	1.0 ± 0.1	–
	0.2 s	1.8 ± 0.1	80 ± 20
	2 s	1.7 ± 0.1	70 ± 20
	8 s	1.7 ± 0.1	70 ± 20
WT (fast)	Dark state	1.00 ± 0.04	–
	0.02 s	1.59 ± 0.03	59 ± 5
	0.3 s	1.56 ± 0.02	56 ± 4
Data WT (not shown) Set 2	Dark state	1.0 ± 0.2	–
	8 s	1.3 ± 0.1	30 ± 28
C71S[b]	Dark state	1.0 ± 0.2	–
	8 s	1.0 ± 0.2	–

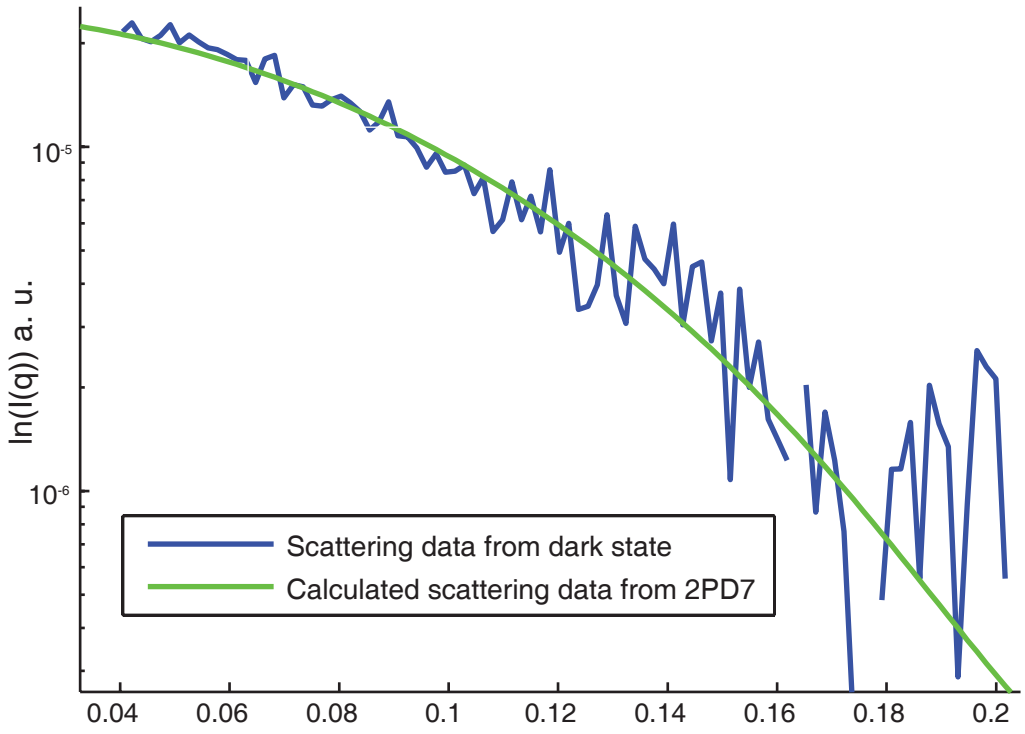

Figure 9.2: Comparison of scattering data to known crystal structure of the VVD monomer. Within the noise, the data is in good agreement with the published crystal structure. The chi-squared value calculated by Crysol[3] is 2.3[2]

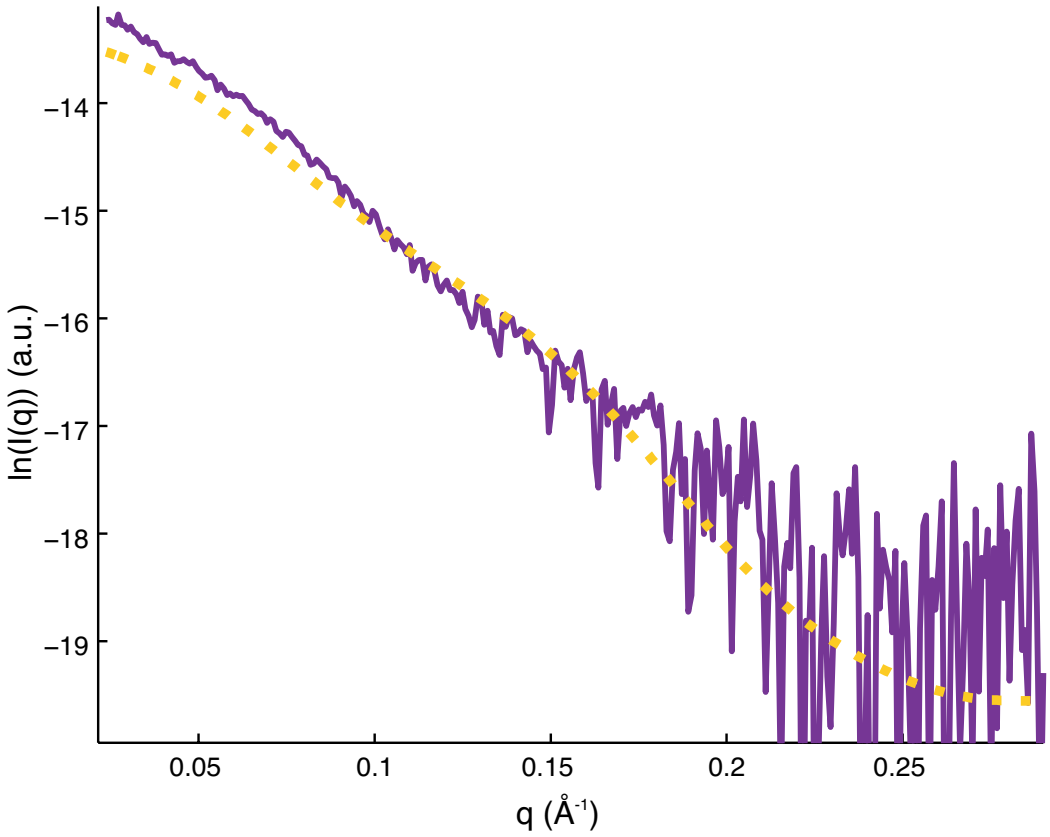

Figure 9.3: Comparison of dimer scattering from data versus crystal structure. The scattering from the light-activated dimer, extracted from the minimization method described in the text, is shown in violet. This is compared to the calculated scatter from the crystallized VVD dimer in yellow. The shape of the two is different, adding to evidence already found by Zoltowski and Crane [4] that the actual light-activated dimer molecule is somewhat different than what is seen in the crystal structure[5].

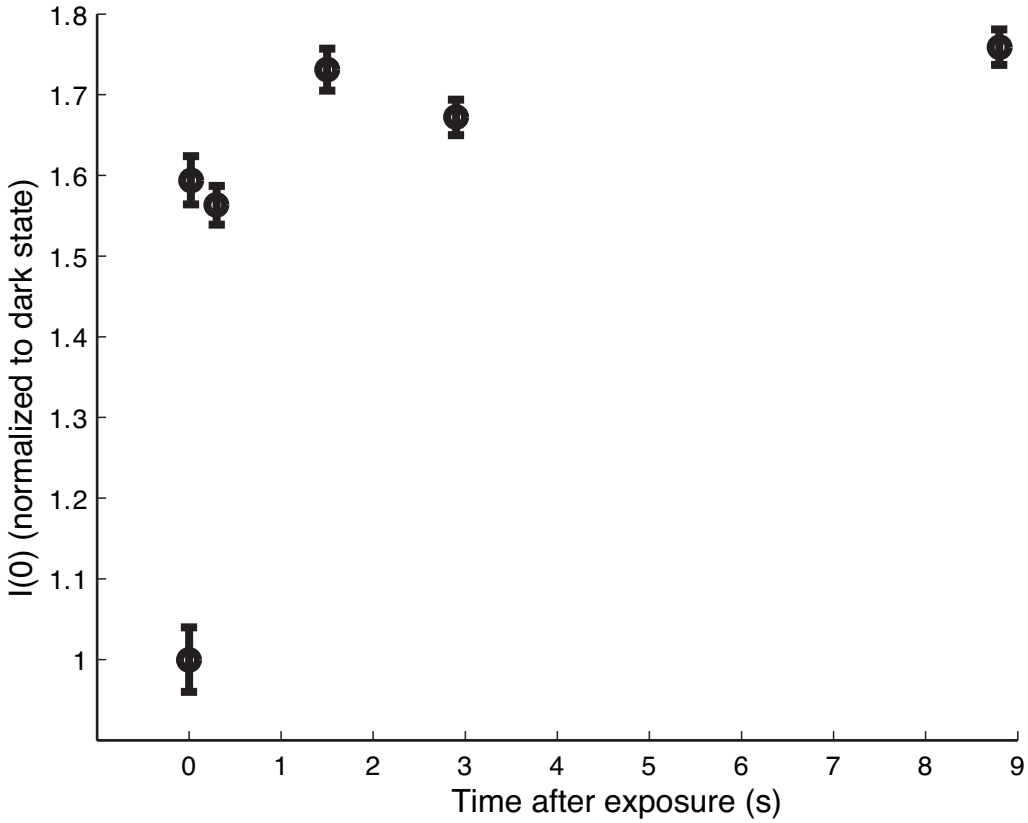

Figure 9.4: $I(0)$ from Guinier analysis from the $\text{Dark}_{\text{crys}}$ wildtype VVD. The value of $I(0)$ has been normalized at t=0 to demonstrate the extend of dimerization in latter time points. There is a slight trend towards larger values at longer times, consistent with a small amount of ongoing dimerization over the range of measurement times. More definitive is the large increase in both $I(0)$ between the dark state (time equals 0) and first light excited point. This demonstrates most molecules in the population form the light-activated dimer before 0.02s[5].

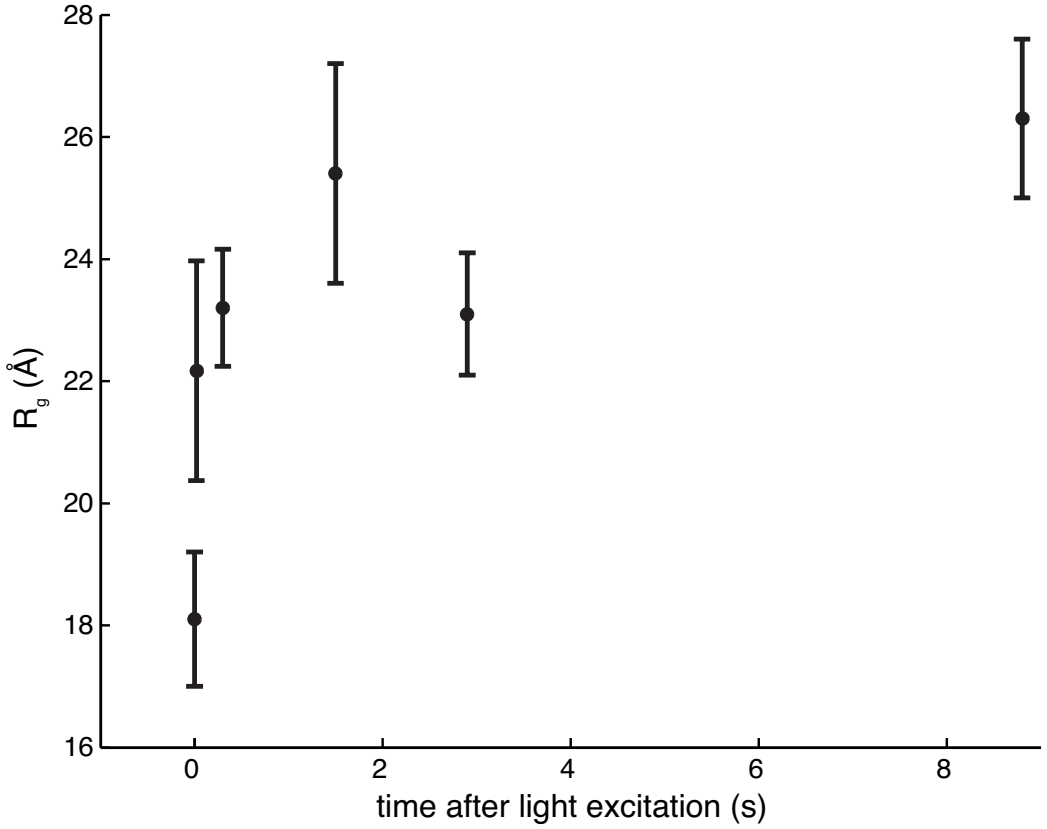

Figure 9.5: R_g versus time after excitation for wildtype VVD with the crystal-like monomer state[5].

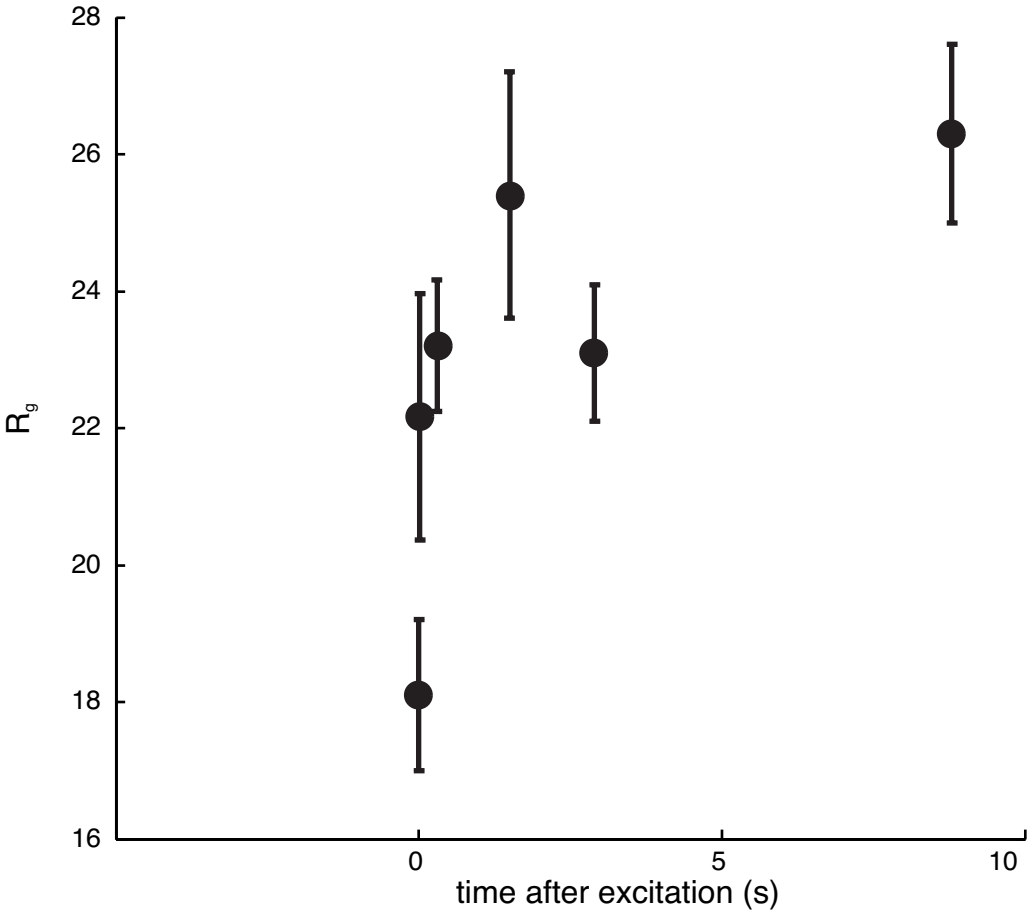

Figure 9.6: R_g versus time after excitation for wildtype VVD with the "alternate" monomer state[5].

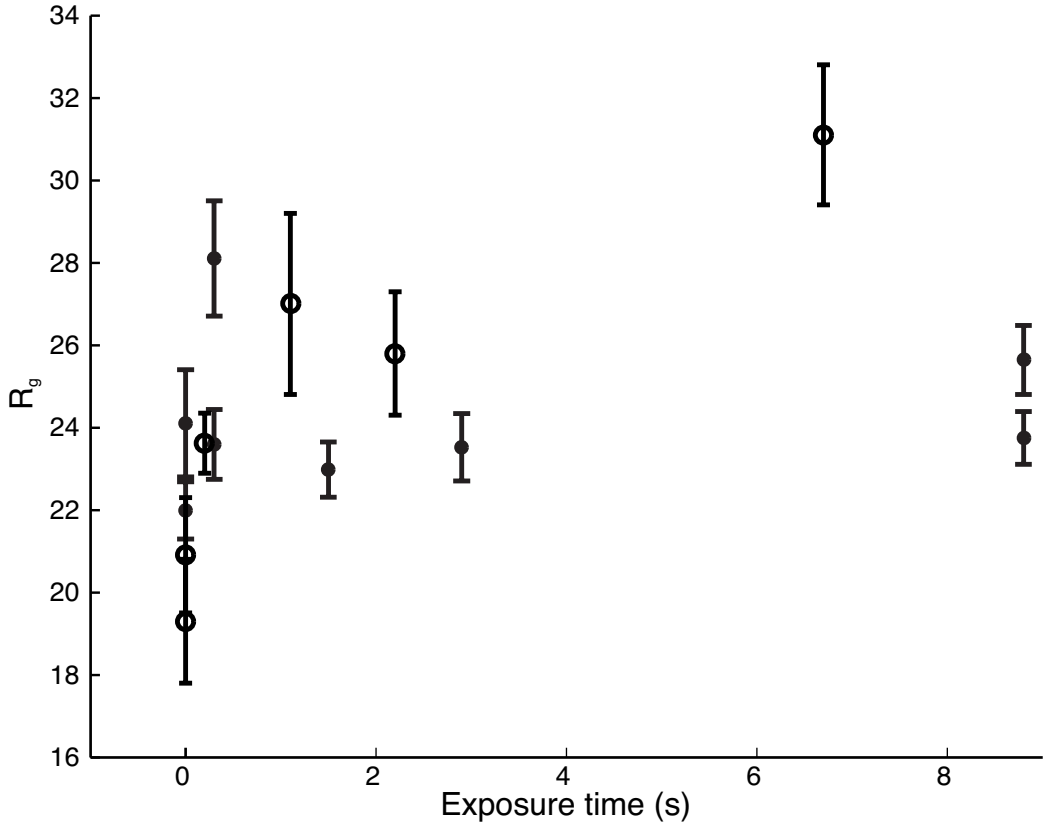

Figure 9.7: R_g versus time after excitation for the C71V:C173S variant. Hollow circles represent data taken with a sheath flow[5].

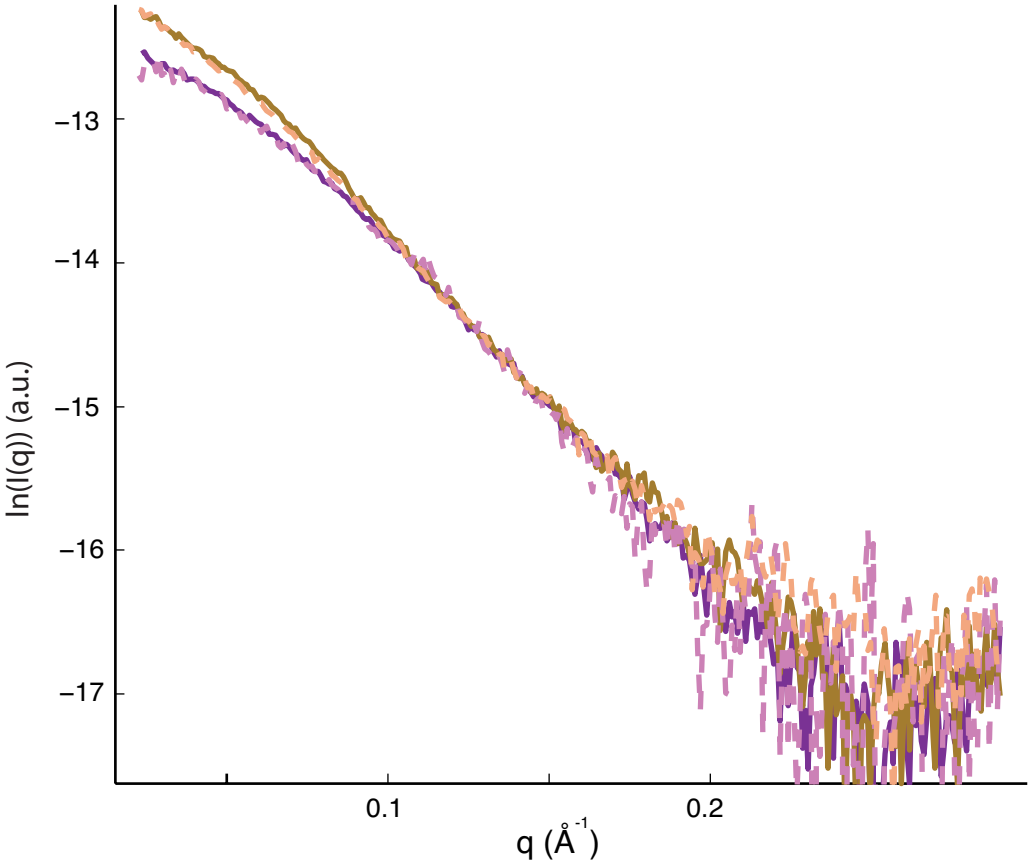

Figure 9.8: Scattering curves of the variant comparing the dark state (lower plots) and the 8.8 s light state collected 24 hours apart. Data from different days plotted on the same axes have been scaled to emphasize the similarity at high q.To ensure that the presence of dimers was not the result of accidental light exposure, we remeasured the scattering profiles after the protein was kept in dark storage for 24 hours. No change was seen in the profiles, confirming the apparent dimerization did not result from an accidental exposure to light. We conclude that this variant dimerizes spontaneously. This together with the fact that the monomer conformation is apparently extended, like the Dark_{alt} structure, implies the extension somehow assists dimerization[5].

REFERENCES

1. Lamb, J. S., L. W. Kwok, X. Qiu, K. Andresen, H. Park, and L. Pollack. "Reconstructing three dimensional shape envelopes from time resolved small angle X-ray scattering data." *J. Appl. Cryst.*, in press.

2. Lamb, J., B. D. Zoltowski, S. A. Pabit, B. R. Crane, and L. Pollack. "Time-resolved dimerization of a PAS-LOV protein measured with photocoupled small angle X-ray scattering." *J. Am. Chem. Soc.*, 2008. Doi:10.1021/ja804236f.

3. Svergun, D., C. Barberato, and M. H. J. Koch. "CRYSOL - A program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates." *J. Appl. Cryst.*, **28**: 768–773, 1995.

4. Zoltowski, B. and B. Crane. "Light activation of the LOV protein vivid generates a rapidly exchanging dimer." *Biochemistry*, **47**(27): 7012–7019, 2008.

5. Lamb, J. S., B. D. Zoltowski, S. A. Pabit, L. Li, B. R. Crane, and L. Pollack. "Illuminating conformational changes of a LOV domain protein with photocoupled small angle X-ray scattering." in preparation.

APPENDIX A

USAGE OF THE BIOLOGIC SFM-400 AT THE CHESS G1 STATION

A.1 Sample cuvette

Reactions initiated by the stopped-flow mixer are typically measured with optical setups. For this reason, cuvettes available for use with it are made of thick quartz. This is inappropriate for SAXS, so we designed alternate sample cells. They are machined from stainless steel and designed to sit in the cuvette holder for the TC-100 style cuvettes. They are 2 mm thick, ideal for work at 10 keV and have either a vertical or horizontal window geometry.

We attach nitride windows using RTV, which is a relatively soft epoxy. We originally tried to use Stycast, but the hardness strained the nitride windows and they broke after time. Nitride windows will inevitably need replacing. Following is the procedure for doing so:

You will need:

- RTV
- your favorite epoxy applicator
- stainless steel cuvette with old epoxy removed
- a microscope slide
- foamy double-sided tape

1. Set the oven to about 90°.

2. Place double-sided tape in the middle of the microscope slide.
3. Carefully apply epoxy in a thin layer to one side of the cuvette.
4. Place cuvette, epoxy side up, on double sided tape.
5. Spin slide/cuvette ensemble. 3 minutes at 300 RPM should be sufficient.
6. Bake for about 5 minutes, so epoxy is not set, but less runny
7. Place first nitride window
8. Bake 10 minutes
9. Repeat procedure on other side, but this time bake for an hour

A.2 Timing

In order to correctly use the stopped-flow mixer, one must carefully consider how to control the timing interval between the mixing and the X-ray exposure. We have used the SFM-400 (Biologic) at Dline and Gline, with the Medoptics detector (both stations) and the Flicam (Gline only). I will outline the procedure for the most recent setup of the mixer with the Flicam at G1. Much of this will be generally applicable, but every detector uses different software, so some aspects of the timing control might need to change if an alternate detector is used.

The SFM is controlled via a software program called MPS v. 2.05. This program allows one to times for each stage of the mixing, and has a call to SYNC which gives a TTL signal. It can also be triggered via a TTL signal, though oddly, if the input signal is high when SYNC is called the output will not be triggered. We have a gray box that converts the input enable signal into a pulse for this reason. Also strangely, there is a delay between triggering and actuating

the program - this is 1.1 s, unless you have the program do no mixing, in which case it is 0.1 s. Typically, we have controlled the timing by using spec to initiate the mixer, and then the mixer to open the shutter via the SYNC terminal. There is a newer version of the SFM software, but it does not have the same SYNC functionality, and does not seem to meet our needs.

In this setup we want to be able to control the shutter in four modes from the spec command line without changing the setup: open/close without collecting and image, collect a dark image, collect an X-ray image, and collect a timed X-ray image after initiating mixing. We would also like accurate flux counts by the PIN diode and ion chambers, which are directly controlled by spec's counter. This means we need to use the signal from spec's counter to open the shutter, and coordinate it with the mixing time and opening the shutter. The counter is controlled partly in software and partly in hardware. The hardware I/O interface is shown in figure A.1. There are three lemo-style connections we are concerned with. ARM IN/ARM OUT are wired together and have a T output to connect to. The GATE input can also be wired. The ARM connector becomes high when the spec computer calls for a count, but if GATE is low, the counting is inhibited. If GATE is high but goes low, the counting will stop, and if it goes high again counting will continue. We use this function to hand off timing control between the SFM and spec computer.

SPEC uses a number of digital motors to control the detector/shutter. These are ENABLE, CCD_GATE, DARK, and SHUT. There are also a number of logic operation boxes on the beamline, shown in Fig. A.1. Finally, operation of the X-ray shutter is controlled via TTL logic. Wiring of these components to achieve the functionality described above is shown in Fig. A.2.

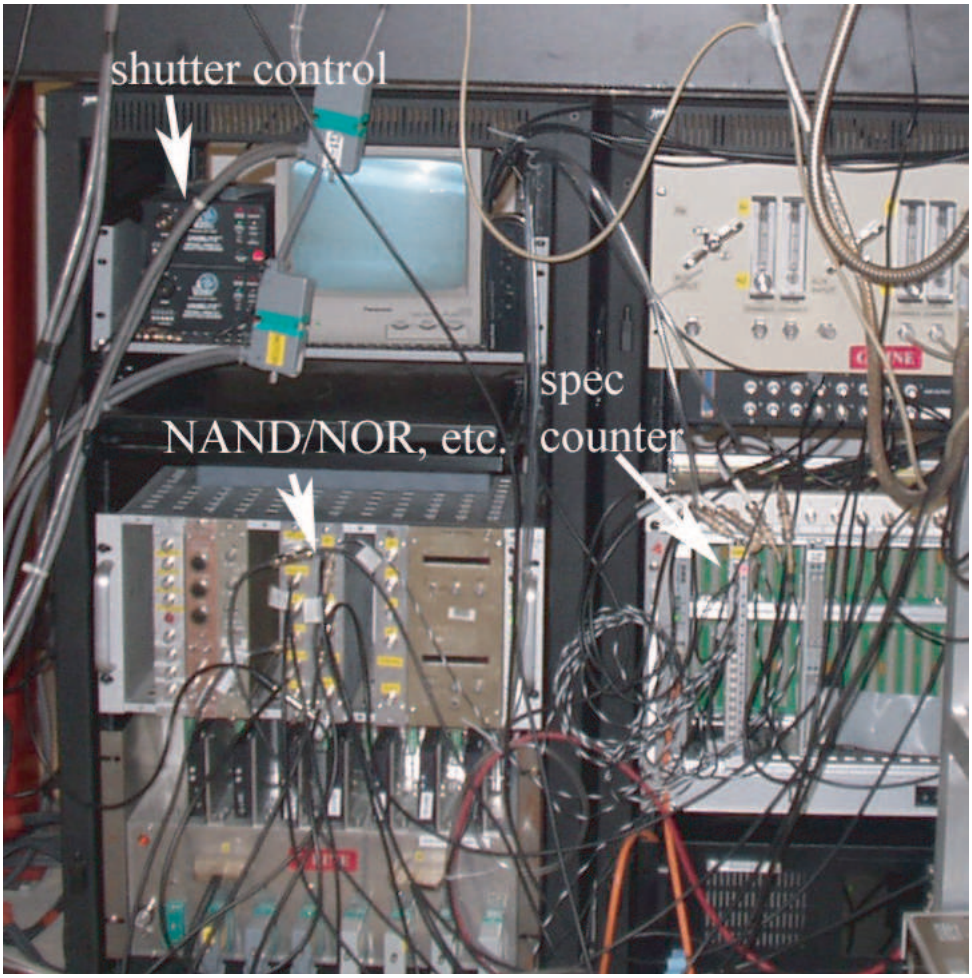

Figure A.1: This figure shows the location of several I/O and logic boxes described in this document. They must be wired correctly to achieve correct timing function.

To take an exposure with the SFM working, the command "mseries delay expt" is called from the spec command line. This sets the spec internal variable MIXER_MODE to 1, which sends in to the correct routines for collecting an exposure. "Delay" should be the *entire* length of the SFM program - the time required for mixing, the exposure time, and any additional delays programmed to make the timing work. "Expt" is the desired X-ray exposure time. Spec moves ENABLE and CCD GATE high to start the mixing sequence and calculates the appropriate delay before sending the command to FLICAM to collect the image.

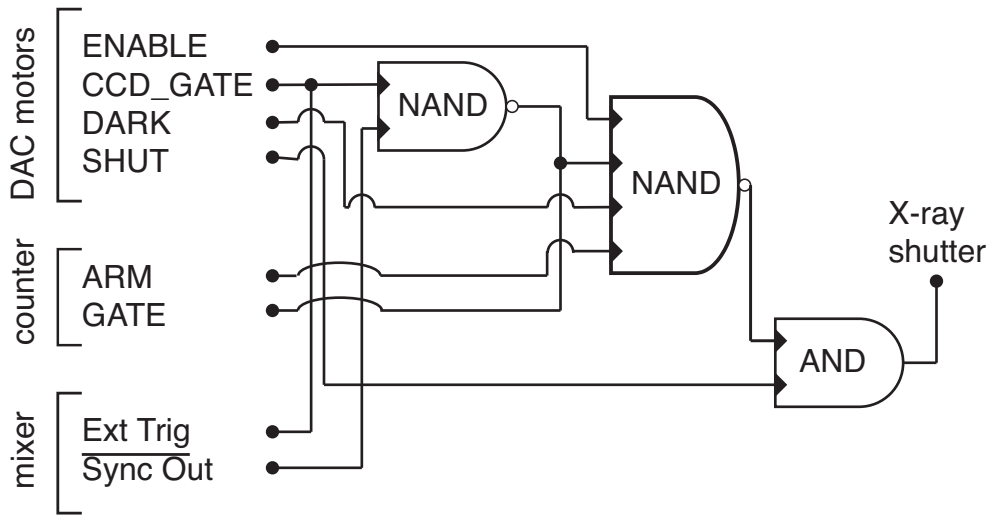

Figure A.2: This schematic shows the wiring connections necessary to operate the SFM at Gline.

Because the amount of time it takes the both spec and the CCD controller computer to process commands is not precise, it is difficult to time the mixing and image collection correctly through software alone. We make up for small discrepancies by "bracketing" the CCD exposure. In other words, the time the detector collects the images is longer than the time for which the x-ray shutter is open. Thus, rather than matching two 100 ms events to withing a ms, we only must place a 100 ms event within a 150 ms interval. In the specific case of the FLICAM the detector automatically integrates for 58 $\pm$ 10 ms longer than instructed, so we make use of this excess to bracket the shutter exposure. There is also a 1.795 $\pm$ 0.005 s flush time which occurs between calling the detector and starting to collect the image. These times are all accounted for in the current programming of mseries, but if a different detector is used, they will need to be investigated and reprogrammed.

Figures A.3 and A.4 show the different timing sequences. To clarify, following is a verbal list of the events that take place when "mseries" is called.

1. Command sent. ENABLE goes high to start the exposure hardware sequence and CCD GATE goes high to trigger the mixer.
2. CCD GATE going high also sends the gate for the spec counter low, preventing the shutter from opening.
3. A few ms before the mixer reaches the last part of its program, the detector begins to integrate
4. The mixer program sends the $\overline{SYNC\,OUT}$ low. This in turn sends the gate for the spec counter high. The shutter opens and spec starts counting.
5. When spec finishes counting, the ARM goes down and the shutter closes. Note this part is controlled by the spec counter, not the SFM. If $\overline{SYNC\,OUT}$ were to change early, spec would never finish counting. It doesn't matter if $\overline{SYNC\,OUT}$ changes late.
6. Shortly after this the integration stops
7. Shortly after this, spec and the detector computer finish any postprocessing, and all the signals return to their default levels

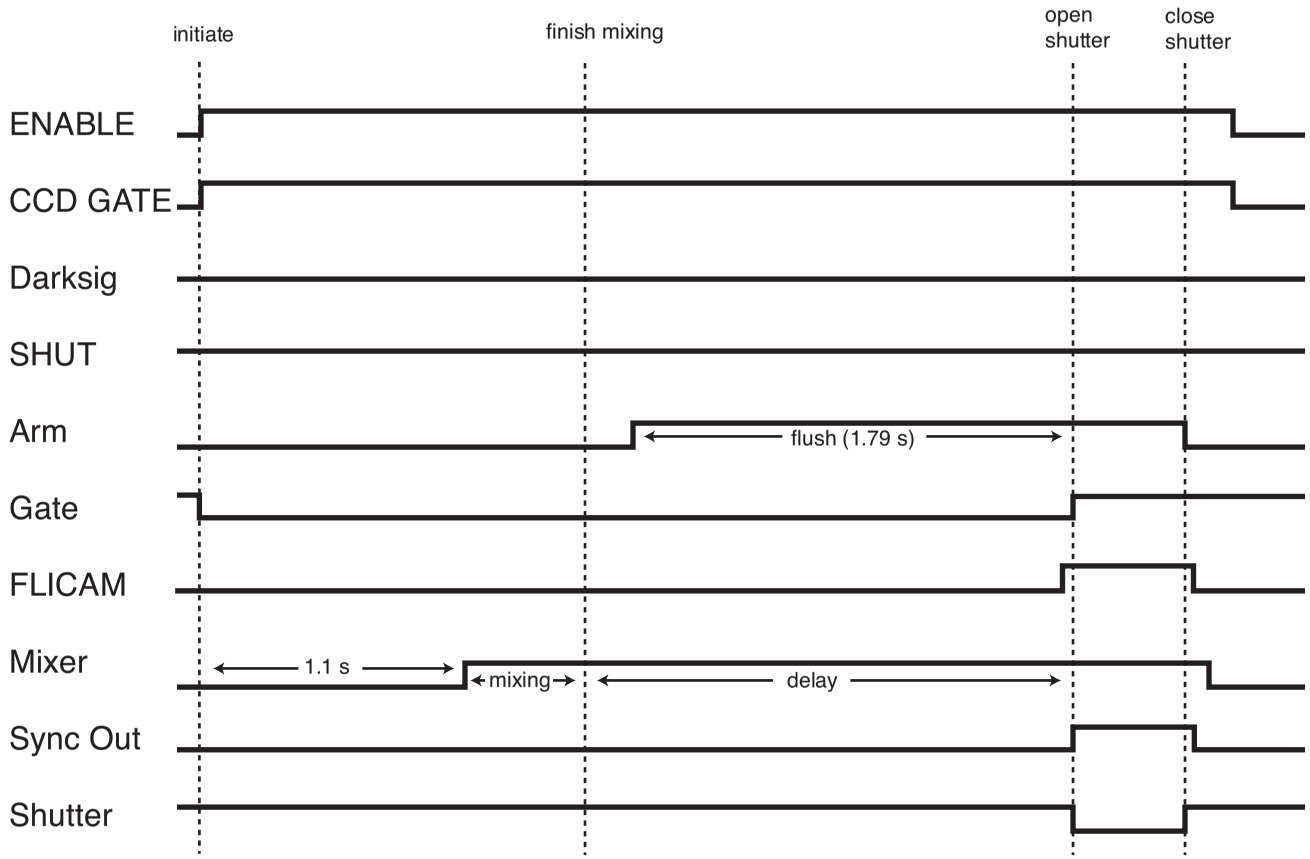

Figure A.3: The timing sequence for the mseries command. Logic components depict their TTL value. Other components, such as the mixer and the detector, are shown as high when they are active.

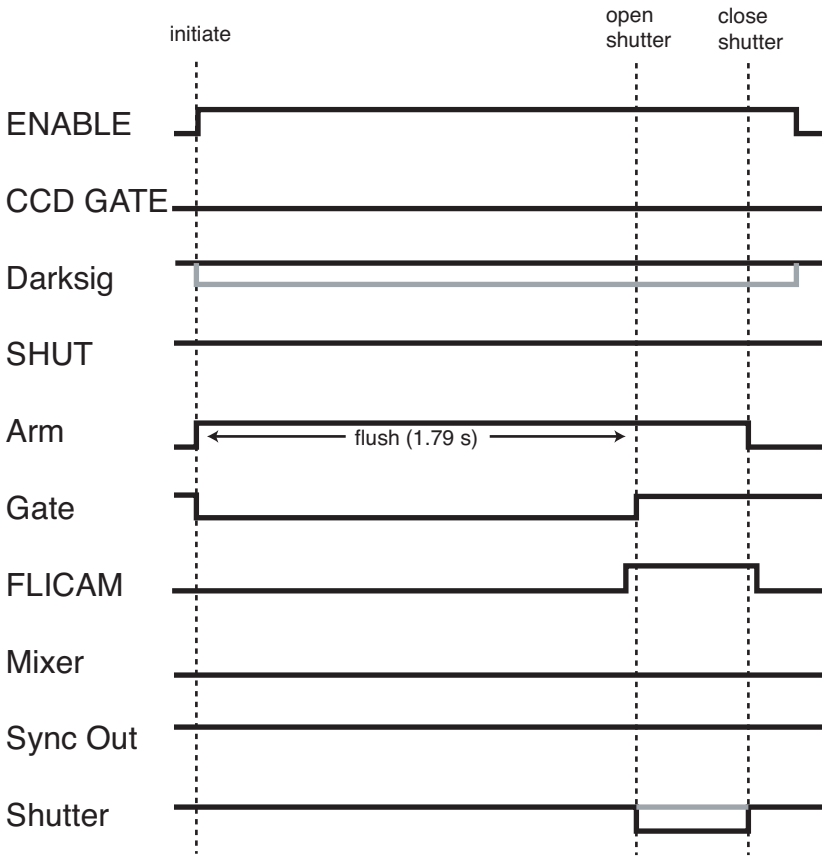

Figure A.4: The timing sequence for the tseries command, with the dseries also shown in gray.

APPENDIX B

INSTRUCTIONS FOR ASSEMBLING THE CONCENTRIC FLOW CELL

List of parts:

- Aluminum flow cell body (Machining diagrams at the end of this appendix)
- Teflon sample inlet port (Machining diagrams at the end of this appendix)
- 2 4 mm stainless steel rods (Thorlabs, NJ) (Other rods will be needed for mounting)
- 1 mm PET tubing (Advanced Polymers, VT) (part # 040050CST)
- 30 G hypodermic tubing (McMaster Carr or Smallparts)
- 3 assembled gripper fittings for 1/16" tubing (Omnifit)
- 4 3-48 spring-loaded ball bearing screws
- 4 3-48 cone-tipped screws
- MA 300 epoxy (Plexus)
- 5 minute epoxy
- Size 002 buna o-ring
- Vacuum grease

Instructions:

1. Fit the aluminum cell body pieces onto the stainless steel rods from Thorlabs and measure the PET tube so it is just a bit longer than the space between the two body pieces. Cut the tubing so it is approximately 0.01"

longer than the spacing. Remove one body piece, place the tubing in the other, and reassemble carefully, avoiding kinking the tubing.

2. Using 5 minute epoxy, glue the tubing in place for the purpose of creating a liquid seal. Do not use the MA 300 epoxy since this is heat shrink tubing and that has a very exothermic cure.

3. Obtain a piece of hypodermic tubing approximately 5/8" long. I found that cutting the tubing with wire cutters and bending the shape back with needle nose pliers and sanding gave cleaner results than any type of saw. However, I have since found out Smallparts will cut such tubing to specific lengths, which would probably be an improvement.

4. Position the piece of hypodermic tubing in the teflon inlet port (I like to use a soldering clamp with 2 alligator clips) and glue in place using the MA 300 epoxy. Make the tubing as concentric to the shape of the inlet port as possible.

5. After the epoxies have thoroughly cured, place the set screws in the 8 threaded holes at the inlet of the aluminum body. Be sure a spring-loaded screw is opposite a cone point. Lightly grease an 002 oring (this helps with the liquid seal) and thread it over the hypodermic tubing.

6. While pushing the teflon port hard into the aluminum body, tighten the set screws. If you do not push at this point, the seal is not likely to be liquid tight.

7. Adjust the set screws so the hypodermic tubing is concentric with the PET tube.

8. Add tubing to the three 1/4-28 ports via the gripper fittings. Be careful not to twist the teflon port relative to the aluminum body. Screw the fittings

in so they're tight, but overtightening is unnecessary and can damage the soft teflon fittings.

9. Test the device for leaks and concentric flow. Many problems can be corrected by adjusting the set screws and reseating the inlet port. If this is insufficient, you may need to reglue. The device can be disassembled mechanically, and remaining epoxy can be scraped off with a razor blade or scalpel.

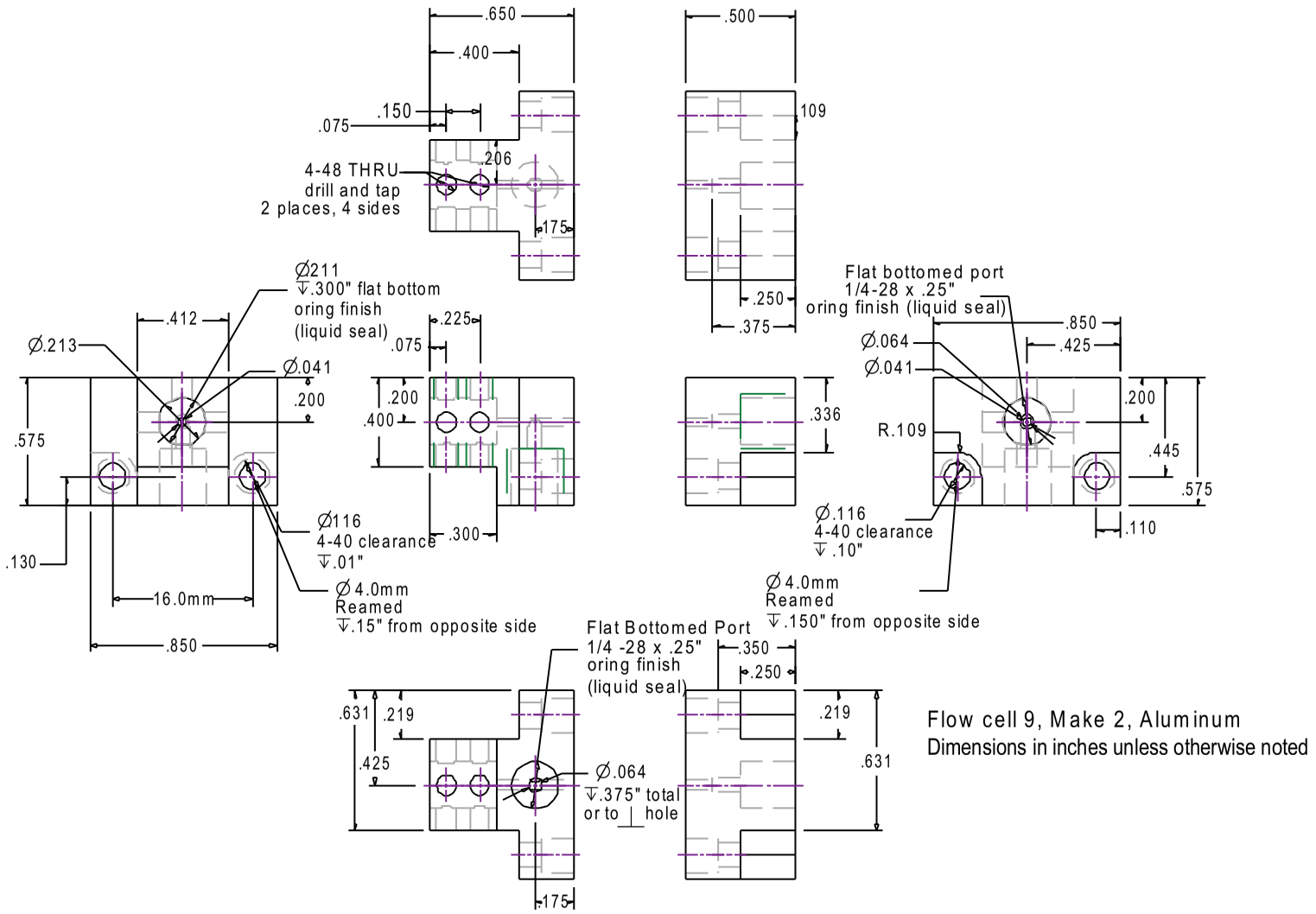

Figure B.1: Flow cell machining diagram. Material is aluminum, units are in inches.

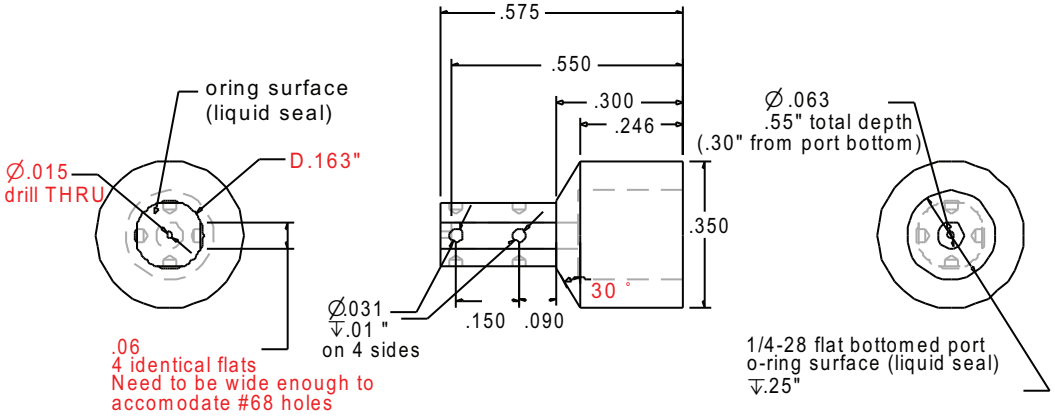

Figure B.2: Inlet port machining diagram, for use with the concentric flow PET cell. Material for this port is teflon.